

LANCS-TH 94-15  
 SUSSEX-AST 94/8-2  
 astro-ph/9408066  
 (August 1994)

# Interpreting Large Scale Structure Observations

David H. Lyth<sup>†</sup> and Andrew R. Liddle<sup>\*</sup>

<sup>†</sup>*School of Physics and Materials,  
 University of Lancaster,  
 Lancaster LA1 4YB. U. K.*

<sup>\*</sup>*Astronomy Centre,  
 Division of Physics and Astronomy,  
 University of Sussex,  
 Brighton BN1 9QH. U. K.*

To appear in the proceedings of  
 “Journée Cosmologie”, held in Paris, June 1994.  
 Presented by D. H. Lyth

## Abstract

The standard model of large scale structure is considered, in which the structure originates as a Gaussian adiabatic density perturbation with a nearly scale invariant spectrum. The basic theoretical tool of cosmological perturbation theory is described, as well as the possible origin of the density perturbation as a vacuum fluctuation during inflation. Then, after normalising the spectrum to fit the cosmic microwave background anisotropy measured by COBE, some versions of the standard model are compared with a variety of data coming from observations of galaxies and galaxy clusters. The recent COBE analysis of Górski and collaborators is used, which gives a significantly higher normalization than earlier ones. The comparison with galaxy and cluster data is done using linear theory, supplemented by the Press-Schechter formula when discussing object abundances of rich clusters and of damped Lyman alpha systems. By focussing on the smoothed density contrast as a function of scale, the observational data can be conveniently illustrated on a single figure, facilitating easy comparison with theory. The spectral index is constrained to  $0.6 < n < 1.1$ , and in particle physics motivated models that predict significant gravitational waves the lower limit is tightened to 0.8.

## 1 Introduction

By ‘large scale structure’ one means galaxies and clusters, with emphasis on their spatial distribution and motions, and also the cosmic microwave background (cmb) anisotropy. We are at present on the verge of a quantum leap in our understanding of large scale structure, because the cmb anisotropy is being measured for the first time [1, 2, 3, 4]. In particular, we may in the foreseeable future verify or rule out the standard model of structure formation, according to which large scale structure arises from a Gaussian, adiabatic density perturbation that is nearly scale invariant at horizon entry. If that model is verified there will be the dazzling prospect of a window on the fundamental interactions on scales approaching

the Planck scale, because the vacuum fluctuation during inflation generates just such a perturbation.

This article has two objectives. One is to equip the non-specialist with a starting point, from which to follow the saga that will unfold during the coming years. To this end we include an extensive discussion of linear perturbation theory, which allows one to translate theoretical input coming from say a model of inflation into a form amenable for comparison with observations. Our other aim is to provide a critical assessment of the present position, in the light of all relevant observations, including the abundance of damped Lyman alpha systems and the latest analysis of the crucial COBE data. These and the other relevant observations are summarized on a single plot. On the same plot are drawn the canonical version of the standard model, and variants which alter the Hubble constant  $h$ , the spectral index  $n$ , and the fraction  $\Omega_\nu$  of any hot dark matter. We discuss the observational constraints on these three parameters, and in particular the constraint on  $n$ , which is of great interest because it can be a sensitive discriminator between inflationary models. As we have emphasised several times in earlier publications [5, 6, 7, 8, 9, 10] a powerful constraint on  $n$  is provided by the long ‘lever arm’ between the very large scale explored by the large angle cmb anisotropy and the smaller scales explored by galaxy and cluster data. On the basis of the presently available data we find that  $0.6 < n < 1.1$ , with lower limit tightened to 0.8 in particle physics motivated models of inflation that generate significant gravitational waves. The lower limit in particular is rather firm because several different types of observation confirm it.

### Three possible models for large scale structure

Three possible models of large scale structure are commonly entertained. The standard model is that it originates as a Gaussian adiabatic density perturbation, whose spectrum is nearly scale independent at horizon entry. This model has been explored far more thoroughly than the other two, and is the only one that will be considered here. An alternative is that large scale structure originates from topological defects, such as cosmic strings. In both of these models the underlying scale invariance means that galaxy and cluster formation is directly related to the magnitude of the large scale cmb anisotropy. It also means that the first objects to form are the progenitors of galaxies. Finally there is the possibility that large scale structure originates from a density perturbation (either adiabatic or isocurvature) with a spectrum that is not even approximately scale independent. In that case structure formation could be very different, with perhaps much lighter objects forming first, and it would not be directly related to the magnitude of the large scale cmb anisotropy.

The standard model has two features which distinguish it from the others, and make it so attractive. One of them concerns theory. If, as is widely supposed, the initial conditions for the hot big bang are set by inflation, then an adiabatic, Gaussian, more or less scale invariant density perturbation is *predicted*. What could be more natural than to suppose that perturbation will explain large scale structure? If that turns out to be so, large scale structure will provide us with a unique window on the nature of the fundamental interactions, because both the magnitude and precise scale dependence of the perturbation are highly model dependent [6, 11, 12, 8].

The other feature concerns phenomenology. The model has been intensively studied and is relatively simple, so that by now one knows how to estimate its predictions for most available types of data. To a first approximation the predictions depend on a single number, specifying the magnitude of the density perturbation at horizon entry. If that number is chosen to fit the cmb anisotropy, all other data can certainly be explained to within a factor of two or three! Of course the burning question is whether the data can actually be explained within their observational uncertainties, which in the best cases are only tens of percent. The answer to that question depends what other parameters are available in the

standard model, and it will be our main focus.

## An overview of the standard model

The standard model assumes that large scale structure originates as an adiabatic, Gaussian density perturbation whose spectrum is more or less scale-independent at horizon entry. The scale dependence (if any) is parameterised as a power law, with a spectral index that by convention is defined so that  $n = 1$  corresponds to scale invariance.

In order to have any chance of agreeing with observation, the standard model requires non-baryonic dark matter, which is more or less cold, and which has a density dominating the baryon density. By ‘cold’ one means that the constituent particles are stable, non-interacting and non-relativistic, at all relevant epochs.

The simplest version of the standard model assumes that  $n = 1$ , and that the (non-baryonic) dark matter is completely cold. It also assumes that the energy density of the universe is critical,  $\Omega = 1$ , with no cosmological constant or other exotic contribution so that there is the standard cosmology with matter domination at present. This critical density,  $n = 1$  CDM model was the favoured one for many years. It contains only two free parameters, which are the normalisation of the spectrum, and the value of the Hubble constant  $H_0 \equiv 100h \text{ km sec}^{-1} \text{ Mpc}^{-1}$ . According to observations having nothing to do with large scale structure (namely, direct observations and measurements of the age of the universe),  $0.4 \lesssim h \lesssim 0.6$ .<sup>1</sup> Assuming the central value  $h = 0.5$  we arrive at a canonical version of the CDM model, whose only free parameter is the magnitude of the scale invariant density perturbation.

The canonical CDM model does not agree with observation because the predicted spectrum of the density perturbation has the wrong scale dependence. On large scales the spectrum is accurately determined by the COBE measurement of the cmb anisotropy, and as one goes down in scale through the regime explored by data on galaxies and clusters it becomes progressively too big compared with the data. How can we reduce the small scale power?

The simplest possibility is to reduce  $h$  below the canonical value  $h = 0.5$ , which reduces the small scale density perturbation because it delays matter domination giving the density perturbation less time to grow. It has recently been noted [13] that this ‘old universe’ option might work if  $h$  is as low as 0.3, but such a value is difficult to reconcile with measurements of  $h$  from Hubble’s law.

Another possibility is to reduce  $n$  below the canonical value  $n = 1$ . This ‘tilted spectrum’ option has been widely investigated [5, 6, 14, 15, 16, 7, 9, 10], and it might be viable with  $n \simeq 0.7$ . Unfortunately most inflation models with tilt also generate significant gravitational waves which ruin this concordance, as we discuss later. Of course one can combine the old universe and tilted spectrum options.

A third option is to change the hypothesis of completely cold dark matter, which reduces the small scale density perturbation because cold dark matter maximises its rate of growth. An attractive possibility, from both a theoretical and observational viewpoint, is to invoke a fraction  $\Omega_\nu$  of hot dark matter in the form of massive neutrinos. This *mixed dark matter* (MDM) model has been widely investigated [17, 18, 19, 20, 21, 22, 8, 23, 10], both with and without the option of allowing  $h$  and  $n$  to depart from their canonical values. We shall see that it may be observationally viable with  $\Omega_\nu = 0.15$  or so and the canonical  $h$  and  $n$ . Other possibilities yet to be investigated fully are to replace the CDM by some form of warm dark matter like sterile neutrinos, or to replace it by decaying or self-interacting dark matter.

---

<sup>1</sup>Higher values of  $h$  are permitted only in low density models, especially those featuring a cosmological constant. Given  $h$ , the baryon density is practically fixed by the standard nucleosynthesis relation  $\Omega_B h^2 = 0.013 \pm 0.002$ .

The final possibility is to reduce the matter density, either with [24, 25, 26] or without [27, 25, 26] a cosmological constant to keep the total energy density critical. In practice one takes the non-baryonic dark matter to be completely cold in that case. This *low density CDM model* has been quite widely investigated and with  $h$  and  $n$  at their canonical values it may be observationally viable with  $\Omega_c$  of order 0.5 or so.

All of this assumes that there is no significant gravitational wave contribution to the cmb anisotropy. A contribution up to 50% or so is not ruled out by present data, and is actually predicted by some models of inflation [6, 11, 12]. Of the models that are well motivated from particle physics, those predicting a significant contribution also predict a spectral index  $n < 1$ , and on the large scales explored by COBE they predict that the relative gravitational wave contribution to the mean square cmb anisotropy is  $R \simeq 6(1 - n)$ . In these models the normalization of the density perturbation is therefore reduced by a factor  $[1 + 6(1 - n)]^{-1/2}$ , and the possible existence of this factor should be taken into account when considering tilt in the spectral index. On the other hand, almost all inflation models suggest that if the spectral index is very close to 1 then the gravitational wave contribution will be negligible.

In this article the focus is on the (critical density) MDM model, which of course includes the critical density CDM model as a special case, allowing  $h$  and  $n$  to vary, and keeping in mind the possibility of gravitational waves. The low density CDM model will be mentioned only briefly, and other possibilities not at all.

## 2 Cosmological perturbations

In this section we give a theoretical overview of cosmological perturbation theory, which will later be amplified and compared with observation. We also derive briefly the inflationary predictions for large scale structure, which will provide a unique window on the fundamental interactions if the standard model is verified.

### 2.1 Linear cosmological perturbation theory

The foundation for the standard model of structure formation is linear cosmological perturbation theory, according to which the perturbations satisfy a set of linear partial differential equations as long as they are sufficiently small [28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38]. Using linear theory, one can follow the growth of an enhancement in the mass density until it begins to collapse to form a gravitationally bound structure. One can also give an almost complete description of the cmb anisotropy.

The perturbations decouple into three separate modes, traditionally called scalar, vector and tensor modes. Vector perturbations are associated with the vorticity of the fluid flow, and are presumably negligible since the vorticity decays. Tensor perturbations are, for practical purposes, freely propagating gravitational waves, and so are very easy to handle. They are predicted at some level by inflation, and might affect the cmb anisotropy.

This leaves the scalar perturbations. They are completely defined by giving, for each particle species, the perturbation in the momentum distribution function<sup>2</sup>. From this one can calculate the perturbation in the matter density, pressure and anisotropic stress of each species. (It turns out these in turn determine the metric perturbation, but the latter is not directly observable.)

For a particle species whose collisions are sufficiently frequent to make it a perfect fluid, the full momentum distribution is not needed and one needs only the energy density and pressure. This is a useful approximation for baryons, electrons and photons before photon

---

<sup>2</sup>In principle the perturbations in the momentum distribution functions have vector and tensor modes as well as the scalar one being considered here, but the vector mode is supposed to be absent, and the tensor mode is negligible because gravitational waves have negligible coupling to matter.

decoupling. Alternatively, if a particle species has negligible random motion it is automatically a perfect fluid with negligible pressure, and one then needs only its mass density. This is the case for cold dark matter and for baryons well after decoupling. These perfect fluid approximations provide roughly correct results for both the matter and the radiation, but are not adequate to address the increasingly accurate observations now becoming available.

In order to define a momentum distribution function one needs a set of worldlines, specifying the observers who measure the momentum. Also, to define the perturbation in this or any other quantity one needs to slice spacetime into hypersurfaces so that the quantity can be split up into an average plus a perturbation.<sup>3</sup> Any choice of worldlines and slicing will do, but a convenient one that we adopt here is to use comoving observers and the hypersurfaces orthogonal to them, which are called comoving hypersurfaces. (Comoving observers by definition move with the total energy flow, which means that they measure zero total momentum density.)

### Independent Fourier modes

It is convenient to expand each perturbation as a Fourier series in comoving coordinates, because in linear theory each comoving Fourier mode evolves independently of the others. Comoving Cartesian coordinates  $\mathbf{x}$  are related to physical Cartesian coordinates  $\mathbf{r}$  by  $\mathbf{r} = a\mathbf{x}$ , where  $a(t)$  is the scale factor of the universe. It will be convenient to normalise  $a = 1$  at present. A generic perturbation will be denoted by  $f$ , and the Fourier series is defined in a comoving box much bigger than the observable universe,

$$f(\mathbf{x}) = \sum_{\mathbf{k}} f_{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{x}) . \quad (1)$$

For each Fourier mode the physical wavenumber is  $k/a$  (where  $k = |\mathbf{k}|$ ), and its present value is just  $k$ . At any epoch, a feature in a perturbation  $f$  with size  $aR$  is described by modes with wavenumber of order  $\sim a/k$ .

### Before horizon entry

A crucial epoch for each scale is horizon entry, when the inverse wavenumber  $a/k$  first falls within the Hubble distance  $H^{-1}$ . For critical density, the scale entering the horizon at matter-radiation equality is  $k^{-1} = (20h^{-1})h^{-1}$  Mpc. Smaller scales enter the horizon during radiation domination.

Before horizon entry there is no time for causal processes to operate, and each comoving region of the universe evolves like a separate Friedmann universe [35, 37]. The evolution of the perturbations is therefore very simple, and can be calculated just by comparing the independent evolution of the separate Friedmann universes.

From now on, we specialise to the standard model of structure formation. The standard model states that the initial perturbations are adiabatic and hence determined by the curvature perturbation, and it also states that the curvature perturbation is Gaussian and has a more or less scale independent spectrum.

### The adiabatic initial condition

The adiabatic initial condition states that well before horizon entry the momentum distribution function of each species is a function only of the total energy density, or equivalently of the temperature. For a perfect fluid species it is enough to specify the energy density and pressure, and for a non-relativistic (matter) species the mass density is enough. What

---

<sup>3</sup>The only exception is if the unperturbed quantity (ie., the average) is time independent, in which case the perturbation is independent of the slicing to first order.

functions of total energy density these quantities are depends on the assumed cosmology, the usual assumption being that radiation (extreme relativistic) species are in thermal equilibrium and have practically zero chemical potential.

The adiabatic initial condition determines all the initial perturbations in terms of the energy density perturbation. The perturbations are indeed adiabatic, because the evolution of every quantity along each comoving worldline is that of a Friedmann universe which is adiabatic. (This evolution defines the perturbations, because they all vanish on a hypersurface of constant energy density.) From the relations  $\rho_m \propto a^3$  and  $\rho_r \propto a^4$  for radiation and matter in a Friedmann universe, one learns that the density contrasts of the radiation and matter species are related by

$$\frac{\delta\rho_m}{\rho_m} = \frac{3}{4} \frac{\delta\rho_r}{\rho_r}. \quad (2)$$

In a Friedmann universe the curvature (measured in a comoving distance unit) is constant. One can therefore define a curvature perturbation, which is constant on each scale well before horizon entry [35]. It is conveniently specified by the quantity

$$\mathcal{R} = -H\delta t \quad (3)$$

where  $\delta t$  is the displacement of the comoving hypersurfaces from flat hypersurfaces [7]. On each scale  $\mathcal{R}_{\mathbf{k}}$  is constant well before horizon entry, and it determines the density perturbation  $\delta\rho_{\mathbf{k}}$ . At horizon entry  $\delta\rho_{\mathbf{k}}/\rho \simeq \mathcal{R}_{\mathbf{k}}$ , so one can think of the curvature perturbation as specifying the magnitude of the density contrast at horizon entry.

## Gaussian perturbations

In comparing theory with observation, one is interested only in the stochastic properties of the perturbations. According to the standard model the initial curvature perturbation is Gaussian, which means essentially that its Fourier components have random phases (apart from the reality condition  $f_{\mathbf{k}} = f_{-\mathbf{k}}^*$ ). By virtue of the adiabatic condition, this Gaussian property is bequeathed to all of the perturbations.

All stochastic properties of a Gaussian perturbation are determined by its spectrum, defined as the modulus squared of its Fourier component. To be precise, we define the spectrum  $\mathcal{P}_f$  of a generic perturbation  $f$  as

$$\mathcal{P}_f(k) = 4\pi(Lk/2\pi)^3 \langle |f_{\mathbf{k}}|^2 \rangle \quad (4)$$

where  $L$  is the comoving size of the box used for the Fourier expansion, and  $\langle \rangle$  indicates that  $|f_{\mathbf{k}}|^2$  has been averaged over a small region of  $k$  space to make it smooth. The spectrum is independent of the direction of  $\mathbf{k}$  because no direction in space is preferred. The normalization is chosen so that according to Parseval's theorem the mean square perturbation  $\sigma_f^2$  is

$$\sigma_f^2 = \int_0^\infty \mathcal{P}_f(k) dk/k. \quad (5)$$

Evaluated at random positions, each perturbation has a Gaussian probability distribution with variance  $\sigma_f^2$ . The spectrum  $\mathcal{P}_T(l)$  of the cmb anisotropy is defined in an analogous way, as essentially the modulus squared of its  $l$ th multipole.

## Scale independence of the initial spectrum

The linear evolution gives all of the spectra in terms of the spectrum  $\mathcal{P}_{\mathcal{R}}$  of the initial curvature perturbation. According to the standard model  $\mathcal{P}_{\mathcal{R}}$  is more or less scale dependent, any scale dependence being parameterised by a power law  $\mathcal{P}_{\mathcal{R}} \propto k^{n-1}$ . (The appearance of  $n-1$  instead of  $n$  is a historical accident.) Exact scale independence corresponds to a

spectral index  $n = 1$ , whereas  $n < 1$  corresponds to a ‘tilted’ spectrum which has less power on small scales. One can also contemplate a ‘blue’ spectrum [6, 40, 41],  $n > 1$ , but this possibility is not favoured by either theory or observation.

### The transfer functions

Except on very large scales, the perturbations evolve in a complicated manner after horizon entry, as the particles move about under the combined effect of gravity and particle collisions. By the present epoch however, the situation has again become simple. On cosmologically interesting scales the scale dependence of the matter density perturbation becomes fixed, and so does that of the anisotropy of the cmb. The spectrum of the density contrast is related to  $\mathcal{P}_{\mathcal{R}}$  by a linear transfer function, and so is the spectrum of the cmb anisotropy. The transfer functions depend in general on the value of  $h$ , on the nature and the amount of the non-baryonic dark matter and on the value of the cosmological constant if one is introduced. The transfer function for the cmb anisotropy also depends on the epoch of re-ionisation if it is late.

### Observable scales

Observations of large scale structure probe a range of scales.

The scales probed by the cmb anisotropy are easy to estimate. It originates at the surface of last scattering, whose distance is now close to  $2H_0^{-1} = 6000h^{-1}$  Mpc, and the thickness of the last scattering surface is about  $7h^{-1}$  Mpc. This means that in round figure it probes scales  $10 \text{ Mpc} \lesssim k^{-1} \lesssim 10^4 \text{ Mpc}$ .

Working out the scales that are probed by galaxy and cluster observations is more difficult. An upper limit of order 100 Mpc comes just from the fact that much more distant regions of the universe have yet to be surveyed in detail, but to describe the lower limit we have to describe structure formation.

In viable versions of the standard model, copious structure formation begins only at a redshift of a few, when objects with a very broad range of masses form at more or less the same epoch. In order for an object to be reasonably stable, the baryons within it have to be able to collapse and dissipate their energy, which is thought to be possible only for masses bigger than  $10^6 M_{\odot}$  or so. The resulting objects are thought to become early galaxies. Afterwards there is a bottom-up picture of structure formation, successively more massive objects forming in sequence ending at the present epoch with rich galaxy clusters of mass around  $10^{15} M_{\odot}$ . The lower limit of the mass range to which the bottom-up picture applies is not very well defined but something like  $10^{11} M_{\odot}$  is a reasonable order of magnitude to have in mind [7].

In the bottom-up picture, each object originates as a slightly overdense region in the early universe, the region attracting more matter until it eventually collapses under its own weight. As long as its overdensity is small, the region expands with the universe and is described by linear theory. The Fourier modes describing it have inverse wavenumber  $a/k$  of order its size and in a critical density universe the mass contained in a sphere with radius  $a/k$  is

$$M(k^{-1}) = 1.16 \times 10^{12} h^2 (k^{-1}/1 \text{ Mpc})^3 M_{\odot} . \quad (6)$$

Thus, observations relating to the bottom-up picture of structure formation explore scales in the range  $1 \text{ Mpc} \lesssim k^{-1} \lesssim 10 \text{ Mpc}$ .

What are the relevant observations? Since big clusters formed recently (in the standard model), each of them can probably be identified with one of the first objects that formed in this mass range. The same may be true of smaller clusters and ‘groups’, with masses  $10^{13} M_{\odot} \lesssim M \lesssim 10^{14} M_{\odot}$ . What about smaller masses? Present day galaxies have  $10^6 M_{\odot} \lesssim M \lesssim 10^{12} M_{\odot}$ , but because of merging and other astrophysics they can probably not be

identified with the first objects forming in this mass range. However one also observes quasars and damped Lyman alpha systems out to a redshift of 3 or 4. Their masses are thought to be in roughly the range  $M_{\odot} \sim 10^{11}M_{\odot}$  to  $10^{13}M_{\odot}$ , and they can probably be identified with the first objects forming in this mass range. The formation of the first objects with mass below this range is not directly observed, but observational constraints on the epoch of re-ionisation should provide an indirect probe in the foreseeable future.

The conclusion is that one can observe the formation of the first structure with mass from  $10^{15}M_{\odot}$  down to perhaps  $10^{11}M_{\odot}$ , thereby exploring the density perturbation on scales  $1 \text{ Mpc} \lesssim k^{-1} \lesssim 10 \text{ Mpc}$ . Bigger scales than this can be explored by looking at the spatial distribution and motion of galaxies and clusters. Smaller scales cannot be explored through present observations because one does not understand galaxy evolution, but in the near future the cmb anisotropy may constrain the epoch of re-ionisation and provide an indirect probe on the scale .01 Mpc.

## 2.2 The inflationary prediction for the initial spectrum

The original motivation [42] for scale independence was simply that it is rather natural, particularly in view of the fact that  $\mathcal{P}_{\mathcal{R}}(k)$  is essentially the spectrum of the density contrast at the epoch of horizon entry, when non-trivial evolution begins. If inflation sets the initial conditions, approximate scale dependence is however *predicted* [43, 35, 44, 45, 46]. Let us see briefly how this works.

During inflation the energy density is supposed to be dominated by the scalar field potential  $V(\phi)$ , and  $\Omega$  is driven to 1 so that<sup>4</sup>

$$H^2 \simeq \frac{1}{3} \frac{8\pi}{m_{Pl}^2} V \quad (7)$$

where  $m_{Pl}^2 = G^{-1}$  is the Planck mass and as usual we are setting  $c = \hbar = 1$ . The spectral index depends only on the first and second derivatives of  $V$  which are conveniently defined by the dimensionless parameters

$$\epsilon \equiv \frac{m_{Pl}^2}{16\pi} \left( \frac{V'}{V} \right)^2 \quad (8)$$

$$\eta \equiv \frac{m_{Pl}^2}{8\pi} \frac{V''}{V} \quad (9)$$

where primes denote derivatives with respect to  $\phi$ . Inflation typically occurs only if

$$\epsilon \ll 1, \quad |\eta| \ll 1. \quad (10)$$

This implies that  $H$  is slowly varying (on the Hubble timescale) so that

$$a \propto \exp(Ht). \quad (11)$$

Independently of its initial value,  $\dot{\phi}$  typically settles down to the ‘slow-roll’ approximation

$$3H\dot{\phi} = -V'. \quad (12)$$

On any spatial hypersurface the gradient of  $\delta\phi$  gives the momentum density, so  $\delta\phi$  vanishes on the comoving hypersurfaces. The curvature of these hypersurfaces is defined in terms of their displacement  $\delta t(\mathbf{x})$  from flat hypersurfaces by Eq. (3). On flat hypersurfaces the inflaton field perturbation is therefore  $\delta\phi = -\dot{\phi}\delta t = \dot{\phi}\mathcal{R}/H$ .

In typical models of inflation  $\delta\phi$  has negligible interaction with itself and other fields, so that each Fourier mode evolves independently (consistent with the assumption that linear

---

<sup>4</sup>Einstein gravity is assumed, which can usually be achieved by a suitable definition of the relevant fields.



cosmological perturbation theory applies during inflation). Quantization is straightforward and the mass of  $\delta\phi$  is typically negligible. In order for inflation to be useful cosmologically interesting scales have to start well within the horizon, and the corresponding Fourier modes of every field must be in the vacuum or the corresponding particles would dominate the energy density and spoil inflation [7].<sup>5</sup> The vacuum expectation value of  $|\delta\phi_{\mathbf{k}}|^2$  is therefore given by the usual flat spacetime calculation (the scalar field version of the Casimir effect that is experimentally verified for the electromagnetic field). It can be evolved forward in time using the classical equation of motion that  $\delta\phi_{\mathbf{k}}$  satisfies in the Heisenberg representation, and using Eqs. (10) and (12) one obtains the result [44, 35, 45, 46]

$$\mathcal{P}_\phi^{1/2} \simeq H/2\pi . \quad (13)$$

Since  $H$  and  $\dot{\phi}$  are slowly varying it follows that the curvature spectrum a few Hubble times after horizon exit is given by

$$\mathcal{P}_\mathcal{R}^{1/2}(k) \simeq \frac{H_*^2}{2\pi\dot{\phi}_*} \quad (14)$$

where the star denotes the epoch of horizon exit  $aH = k$ . This value is retained until the scale  $k$  enters the horizon again, after inflation has ended. Using Eq. (7) it can be written

$$\mathcal{P}_\mathcal{R}(k) \simeq \frac{1}{24\pi^2} \left( \frac{8\pi}{m_{Pl}^2} \right)^2 \frac{V_*}{\epsilon_*} . \quad (15)$$

The observed normalization  $\mathcal{P}_\mathcal{R}^{1/2} \simeq 5.8 \times 10^{-5}$  is a powerful constraint on models of inflation. In particular, since  $\epsilon_* \ll 1$  the energy scale  $V_*^{1/4}$  must be less than about  $10^{16}$  GeV and so must the temperature after inflation. This last requirement makes life quite difficult for cosmic string theories of large scale structure though the situation is not hopeless [40].

Using Eqs. (10) and (12) the spectral index  $n = d \log \mathcal{P}_\mathcal{R} / d \log k$  is given by [6, 46]

$$n(k) - 1 \simeq 2\eta_* - 6\epsilon_* . \quad (16)$$

In typical models of inflation  $\eta$  and  $\epsilon$  vary slowly on the Hubble timescale. Since cosmologically interesting scales range over only about four decades of  $k$ , this means that  $n(k)$  can be regarded as a constant leading to the power law  $\mathcal{P}_\mathcal{R} \propto k^{n-1}$ .

Contrary to the view that is sometimes expressed, there is nothing very controversial about this calculation. Provided that  $\delta\phi$  has negligible interactions its vacuum fluctuation well before horizon entry is a standard result of flat spacetime field theory, and its subsequent evolution is given by the classical equation of motion. If Eqs. (10) and (12) are valid the quoted result is obtained, but if these conditions fail one can still calculate the evolution in a straightforward manner using for instance the formalism described in [46]. The predicted spectrum will then generally have a strong departure from scale invariance and the model will be in danger of being ruled out by observation. Even the assumption of negligible interaction can perhaps also be relaxed, leading to a non-Gaussian perturbation [47]. Again, the non-Gaussianity endangers the model.

A similar calculation [48, 7, 46] for the gravitational waves gives their relative contribution to the spectrum of the cmb anisotropy, which on large scales is  $R \simeq 12\epsilon$ . For the particle physics motivated models in which  $R$  is significant,  $V(\phi)$  is well approximated by an exponential or high power-law, which gives  $\eta \simeq 2\epsilon$  and hence  $R \simeq 6(1 - n)$ .

---

<sup>5</sup>We are here making the usual assumption that the curvature is negligible on the present Hubble scale, i.e., that the density is critical counting the contribution of any cosmological constant. The assumption can be relaxed, but then one cannot invoke flat spacetime field theory to define the vacuum. There is however a natural definition, and if it is used and the inflaton field is assumed to be in the vacuum the formulas below still hold [39].

Measured values of  $1 - n$  or  $R$  significantly different from zero would provide a unique window on the inflationary potential, and hence on the fundamental interactions. Even the upper bounds on these quantities that already exist rule out the usual versions of ‘extended inflation’, because [6] such models give both  $n \lesssim 0.7$  and  $R \simeq 6(1 - n)$ .

Now we study in turn the three observable perturbations, which are the matter density perturbation, the peculiar velocity and the cmb anisotropy.

### 3 The Matter Density Perturbation

In the Newtonian regime cosmological perturbation theory is a straightforward application of fluid flow equations, taking into account where necessary particle diffusion and free-streaming (collisionless particle movement). On large scales and in the early universe the Newtonian treatment is inadequate and must be replaced by one based on general relativity, but here too it is possible to derive fluid flow equations using a treatment strongly resembling the Newtonian one [33, 34, 35, 37, 38, 36]. We employ that treatment here instead of the more usual metric perturbation approach [28, 31, 32].<sup>6</sup>

#### 3.1 The fluid flow equations

The strategy is to populate the universe with comoving observers, who define physical quantities in their own region. A crucial concept is the *velocity gradient*  $u_{ij}$ . It is defined by each comoving observer, using locally inertial coordinates in which the observer is instantaneously at rest, as the gradient of the velocity  $u^i$  of nearby comoving observer,

$$u_{ij} \equiv \partial_j u_i . \quad (17)$$

(Note that  $u_i$  is defined locally, with only its gradient having global significance.) The velocity gradient can be uniquely decomposed into an antisymmetric vorticity  $\omega_{ij}$ , a symmetric traceless shear  $\sigma_{ij}$ , and a locally defined Hubble parameter  $H$ ,

$$u_{ij} = H\delta_{ij} + \sigma_{ij} + \omega_{ij} . \quad (18)$$

In the limit of homogeneity and isotropy,  $\sigma_{ij} = \omega_{ij} = 0$ . Then the universe is homogeneous on comoving hypersurfaces (those orthogonal to comoving worldlines).<sup>7</sup> The continuity equation is

$$\frac{d\rho}{d\tau} = -3H(\rho + p) . \quad (19)$$

where  $\rho$  is the energy density,  $p$  is the pressure and  $\tau$  is the proper time measured by comoving observers. The deceleration of gravity is given by

$$\frac{dH}{d\tau} = -H^2 - \frac{4\pi G}{3}(\rho + 3p) . \quad (20)$$

These lead to the Friedmann equation

$$H^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} . \quad (21)$$

The constant  $K$  is a measure of the curvature of comoving hypersurfaces, and we assume critical density  $\Omega = 1$  which corresponds to  $K = 0$ .

---

<sup>6</sup>The same equations can be derived using the ‘gauge invariant’ version of the metric perturbation approach [31, 32], but then the physical significance of the quantities is obscured, and one loses the connection with the Newtonian treatment [36].

<sup>7</sup>As noted in Section 3 this definition has to be modified if the vorticity is not negligible.

Now consider perturbations. One can show from angular momentum conservation that  $\omega_{ij}$  decays like  $(\rho + p)a^{-5}$ . It is therefore presumably negligible though we shall see in the next section how it may be handled if desired. We will also see there how to calculate  $\sigma_{ij}$  which is *not* negligible, but for the moment we need only  $H$ .

On a given comoving hypersurface, each quantity can be split into an average plus a perturbation,

$$\rho(\mathbf{x}, t) = \bar{\rho}(t) + \delta\rho(\mathbf{x}, t) \quad (22)$$

$$p(\mathbf{x}, t) = \bar{p}(t) + \delta p(\mathbf{x}, t) \quad (23)$$

$$H(\mathbf{x}, t) = \bar{H}(t) + \delta H(\mathbf{x}, t) . \quad (24)$$

The time coordinate  $t$  labels the hypersurfaces, and will be taken to be the average of  $\tau$ . We would like to choose the space coordinates  $\mathbf{x} = (x^1, x^2, x^3)$  to be comoving coordinates, related to Cartesian coordinates by  $r^i = ax^i$ , with  $a$  the average scale factor given by  $\dot{a}/a = \bar{H}$ . This cannot be done exactly, because the expansion is not isotropic and the comoving hypersurfaces are not flat. However these effects are of first order, and can therefore be ignored when describing perturbations which are themselves of first order. In other words *all perturbations ‘live’ in the unperturbed Friedmann universe.*

Along each comoving worldline the continuity equation Eq. (19) is unchanged. Ignoring anisotropic stress the Friedmann equation receives extra terms to become [37, 38],

$$\frac{dH}{d\tau} = -H^2 - \frac{4\pi G}{3}(\rho + 3p) - \frac{1}{3} \frac{\nabla^2 \delta p}{\rho + p} + (\text{anisotropic stress term}) . \quad (25)$$

This equation is called the Raychaudhuri equation. The anisotropic stress term is typically of the same order of magnitude as the exhibited one, or smaller.

### 3.2 Separate Friedmann universes

If the additional terms in Eq. (25) are negligible, the evolution along each comoving worldline is identical with that in a Friedmann universe. The Friedmann equation Eq. (21) is obeyed, with some  $K$  that is constant along each comoving worldline. Assuming critical density its average vanishes, so we write  $K \equiv \delta K(\mathbf{x})$  to remind ourselves that it is a perturbation.

To see when the additional terms will be negligible it is enough to compare the exhibited term with the density perturbation  $-4\pi G\delta\rho/3$  coming from the second term (it is straightforward to verify that the latter does not cancel with other terms, and we have just noted that the anisotropic stress term is not expected to dominate the exhibited one). For a given Fourier component  $\nabla^2 = -(k/a)^2$ , and since the density is critical the relative contribution of the exhibited term is  $\sim (k/aH)^2(\delta p_{\mathbf{k}}/\delta\rho_{\mathbf{k}})$ . Well before horizon entry the adiabatic condition ensures that  $\delta p_{\mathbf{k}} \lesssim \delta\rho_{\mathbf{k}}$ , and the relative contribution is small. It is also small sufficiently long after matter domination, because then  $\delta p$  becomes negligible. The conclusion is that *on each scale a comoving region evolves like a Friedman universe before horizon entry, and begins to do so again sufficiently long after horizon entry.*

Given this local Friedmann evolution one can calculate the density perturbation in terms of the curvature perturbation  $\mathcal{R}$  [35]. It can be shown [31, 7] that the definition Eq. (3) is equivalent to<sup>8</sup>

$$\mathcal{R}_{\mathbf{k}} \equiv \frac{3}{2} \frac{\delta K_{\mathbf{k}}}{k^2} . \quad (26)$$

Perturbing the Friedmann equation gives, to first order,

$$2H\delta H_{\mathbf{k}} = \frac{8\pi G}{3}\delta\rho_{\mathbf{k}} - \frac{\delta K_{\mathbf{k}}}{a^2} . \quad (27)$$

---

<sup>8</sup>The factor  $k^2$  means that  $\mathcal{R}_{\mathbf{k}} = (3/2)(\delta K_{\mathbf{k}}/a^2)(a^2/k^2)$  measures the curvature perturbation in units of the relevant scale  $a/k$ . (The intrinsic curvature scalar of the comoving hypersurfaces is  $R^{(3)} = 6K/a^2$ .) Yet another interpretation of  $\mathcal{R}$  is that it is essentially the Newtonian gravitational potential caused by  $\delta\rho$ .

We also need to perturb the continuity equation Eq. (19). In doing so one has to remember that the comoving worldlines are not geodesics because of the pressure gradient. As a result, the proper time interval  $d\tau$  between a pair of comoving hypersurfaces is position dependent. One can show (using essentially the Lorentz transformation between nearby observers) that its variation with position is given by [38, 7]

$$\frac{d\tau}{dt} = \left(1 - \frac{\delta p}{\rho + p}\right). \quad (28)$$

Taking this into account, perturbing the continuity equation gives

$$\frac{d\delta\rho_{\mathbf{k}}}{dt} = -3(\rho + p)\delta H_{\mathbf{k}} - 3H\delta\rho_{\mathbf{k}}. \quad (29)$$

Eliminating  $\delta H_{\mathbf{k}}$  from Eqs. (27) and (29) then gives for the density contrast  $\delta = \delta\rho/\rho$

$$\frac{2H^{-1}}{5+3w} \frac{d}{dt} \left[ \left( \frac{aH}{k} \right)^2 \delta_{\mathbf{k}} \right] + \left( \frac{aH}{k} \right)^2 \delta_{\mathbf{k}} = \frac{2+2w}{5+3w} \mathcal{R}_{\mathbf{k}} \quad (30)$$

where  $w = p/\rho$ . During any era when  $w$  is constant, Eq. (30) has the solution (dropping a decaying mode)

$$\left( \frac{aH}{k} \right)^2 \delta_{\mathbf{k}} = \frac{2+2w}{5+3w} \mathcal{R}_{\mathbf{k}}. \quad (31)$$

In the radiation dominated era before horizon entry this becomes

$$\left( \frac{aH}{k} \right)^2 \delta_{\mathbf{k}} = \frac{4}{9} \mathcal{R}_{\mathbf{k}}(\text{initial}) \quad (32)$$

and in the matter dominated era it becomes

$$\left( \frac{aH}{k} \right)^2 \delta_{\mathbf{k}} = \frac{2}{5} \mathcal{R}_{\mathbf{k}}(\text{final}). \quad (33)$$

As the labels imply, we are regarding the value of  $\mathcal{R}_{\mathbf{k}}$  during the first era as an ‘initial condition’, which determines its value during the ‘final’ matter dominated era.

For future reference note that during matter domination,  $H \propto t^{-1} \propto a^{-3/2}$  and

$$\delta_{\mathbf{k}} \propto a \quad (\text{matter domination}). \quad (34)$$

### The Newtonian picture

Eqs. (21) and (25), which determine the evolution of the energy density perturbation are valid in the framework of general relativity. For the matter dominated case  $p \ll \rho$  they have however the same form as the Newtonian equations, and at a given epoch the Newtonian picture indeed applies to the universe on scales that are sufficiently far inside the horizon [29]. Thus we have demonstrated that the Newtonian equations for the density perturbation are valid during matter domination, even on scales that are too large for the Newtonian picture itself to be valid. We shall shortly demonstrate that the same is true for the equations that describe the peculiar velocity.

### 3.3 The transfer function

On scales  $k^{-1} \gg (20h^{-1})h^{-1} \text{ Mpc}$ , horizon entry is long after matter domination so that the initial and final eras overlap and  $\mathcal{R}_{\mathbf{k}}(\text{initial}) = \mathcal{R}_{\mathbf{k}}(\text{final})$ . On smaller scales there is a *transfer function*  $T(k)$ , which may be defined by

$$\mathcal{R}_{\mathbf{k}}(\text{final}) = T(k) \mathcal{R}_{\mathbf{k}}(\text{initial}). \quad (35)$$

The corresponding density contrast is given by Eq. (33), and it applies to each type of matter [38]. It is Gaussian, and its spectrum is

$$\mathcal{P}_\delta(k) = \frac{4}{25} \mathcal{P}_\mathcal{R}(k) T^2(k) \left( \frac{k}{aH} \right)^4 . \quad (36)$$

The transfer function depends on the nature of the dark matter, and on the value of the Hubble constant. The main physical effects are the following [29, 30, 32]

- The cold dark matter density contrast grows, as overdense regions attract more matter towards them. The growth is slow before matter domination, but rapid thereafter.
- The baryons, electrons and photons form a tightly coupled fluid before decoupling at redshift  $z \sim 1000$ , whose density contrast oscillates as a standing wave. The oscillations are damped because the photons diffuse out of overdense regions, and they carry some of the baryons with them.
- After decoupling the photons travel freely.
- After decoupling, the baryon density contrast grows to match that of the cold dark matter on scales  $\gtrsim 10^6 M_\odot$ , but continues to oscillate for a long time on smaller scales (this may account for the absence of galaxies with  $M \lesssim 10^6 M_\odot$ ).
- Massless neutrinos free stream (travel freely) out of any density enhancement.
- Any hot dark matter (massive neutrinos) free streams while it is relativistic, and then its density contrast grows.

On each scale, the transfer function is applicable after the random particle motion has become negligible. In the absence of hot dark matter this happens soon after decoupling (on scales bigger than  $10^6 M_\odot$ ). In the presence of hot dark matter it happens only after a redshift given by [19]

$$k^{-1} \gtrsim 0.11(1+z)^{1/2} \Omega_\nu \text{ Mpc} . \quad (37)$$

As we are interested only in scales  $\gtrsim 1 \text{ Mpc}$ , the random motion has become negligible by the present. It is still marginally significant at  $z \sim 4$  on the scales relevant for the formation of the quasars and damped Lyman alpha systems, but the effect is small compared with observational uncertainties and we shall not include it.

## The smoothed density contrast

The linear theory described so far applies only to small perturbations. Except perhaps at early times, this means that one has to smooth all quantities on a suitably large scale before using it. Adopting the simple ‘top hat’ prescription, the density contrast  $\delta(\mathbf{x})$  at each point is replaced by its average  $\delta(\mathbf{x}, R)$  within a comoving sphere, whose present radius  $R$  defines the smoothing scale. Instead of specifying  $R$  one can specify the average mass  $M(R)$  within the sphere, which assuming critical matter density  $\Omega_0 = 1$  is given by

$$M(R) = 1.16 \times 10^{12} h^2 (R/1 \text{ Mpc})^3 M_\odot . \quad (38)$$

Top hat smoothing multiplies each Fourier coefficient by the window function

$$W(kR) = 3 \left( \frac{\sin(kR)}{(kR)^3} - \frac{\cos(kR)}{(kR)^2} \right) . \quad (39)$$

It therefore filters out Fourier modes with  $k^{-1} \lesssim R$ . On the other hand, the spectrum  $\mathcal{P}_\delta(k)$  of the unsmoothed density contrast decreases rapidly as  $k^{-1}$  increases on scales  $k^{-1} \gtrsim 1 \text{ Mpc}$ ,

corresponding to the fact that the probability of finding a feature with size  $a/k$  decreases rapidly. As a result, most of the structure in the smoothed density contrast at a given epoch has size of order  $aR$ , for smoothing scales  $R \gtrsim 1$  Mpc.

The mean square  $\sigma^2(R)$  of the smoothed density contrast is

$$\sigma^2(R) = \int_0^\infty W^2(kR) \mathcal{P}_\delta(k) dk/k . \quad (40)$$

At any epoch it can be made arbitrarily small by choosing a sufficiently large filtering scale  $R$ , but it grows with time. As long as  $\sigma(R) \lesssim 1$ , linear theory applies except in those rare regions of space where  $\delta(\mathbf{x}, R) \gtrsim 1$ . These regions correspond to gravitationally bound objects with mass of order  $M(R)$  and linear theory does not apply to them, but since they are rare they will not significantly affect the evolution of  $\sigma(R)$ , or of the Fourier modes with  $k^{-1} \gtrsim R$  that dominate it. When  $\sigma(R)$  grows to become of order 1 on the other hand, the formerly rare gravitationally bound regions become common, and the evolution becomes non-linear on scales  $k^{-1} \sim R$ . Since  $\sigma(R) \propto a \propto (1+z)^{-1}$  (Eq. (34)), this occurs at a redshift given by

$$1 + z_{\text{nl}}(R) = \sigma_0(R) \quad (41)$$

where the 0 denotes the present value of the *linearly evolved* quantity. As we shall see  $z_{\text{nl}}$  is no more than a few in observationally viable versions of the standard model, even on the smallest relevant scales. Before this epoch gravitationally bound objects are rare, and whether they are too rare to account for what is observed is one of the crucial issues that we shall address.

## 4 The peculiar velocity

Associated with the density perturbation is a peculiar velocity field, specifying the departure from uniform Hubble flow that occurs as matter falls into overdense regions.

### 4.1 The Newtonian regime

The Newtonian picture [29] applies to the universe after matter domination on scales sufficiently far inside the horizon, and in particular it applies at present to the sphere of radius a few hundred Mpc around us within which detailed observations of galaxies and clusters have been made .

In the Newtonian picture there is a globally defined fluid velocity field  $\mathbf{u}$ , and choosing the reference frame so that  $\mathbf{u}$  vanishes at the origin the peculiar velocity field  $\mathbf{v}$  is defined<sup>9</sup> by

$$\mathbf{u}(\mathbf{r}) - \mathbf{u}(0) = \bar{H}\mathbf{r} + \mathbf{v}(\mathbf{r}) - \mathbf{v}(0) , \quad (42)$$

An equivalent statement in terms of the velocity gradient Eq. (18) is

$$\delta u_{ij} = \partial_i v_j . \quad (43)$$

where  $\partial_i = a^{-1} \partial / \partial x^i$ . Like any vector field,  $\mathbf{v}$  can be written  $\mathbf{v} = \mathbf{v}^{\text{L}} + \mathbf{v}^{\text{T}}$ , where the transverse part  $\mathbf{v}^{\text{T}}$  satisfies  $\partial_i v_i^{\text{T}} = 0$  and the longitudinal part is of the form  $\mathbf{v}^{\text{L}} = \nabla \psi_v$ . Eq. (18) defines the local Hubble parameter  $H$ , shear  $\sigma_{ij}$  and vorticity  $\omega_{ij}$ , this last being given by

$$\omega_{ij} = \frac{1}{2} (\partial_i v_j^{\text{T}} - \partial_j v_i^{\text{T}}) . \quad (44)$$

Angular momentum conservation gives  $\omega_{ij} \rightarrow a^{-2}$ , so  $\mathbf{v}^{\text{T}}$  decays like  $a^{-1}$  and may be dropped (remember that  $\partial_i \equiv a^{-1} \partial / \partial x^i$ ).

---

<sup>9</sup>Up to a constant, which can be chosen so that the average of  $\mathbf{v}$  vanishes.

Taking  $\mathbf{v}$  to be purely longitudinal, it is determined by the density perturbation in the following way. First, take the trace of Eq. (18) to learn that  $\nabla \cdot \mathbf{v} = 3\delta H$ . From Eqs. (27), (26) and (31) and the matter domination relation  $Ht = 2/3$ , it follows that

$$\nabla \cdot \mathbf{v} = -(4\pi G\delta\rho)t. \quad (45)$$

The solution of this equation is

$$\mathbf{v} = -t\nabla\psi \quad (46)$$

or

$$v_i(\mathbf{x}, t) = -(t/a) \frac{\partial\psi(\mathbf{x})}{\partial x^i} \quad (47)$$

where

$$\psi(\mathbf{x}) = -Ga^{-2} \int \frac{\delta\rho(\mathbf{x}', t)}{|\mathbf{x}' - \mathbf{x}|} d^3x'. \quad (48)$$

The factor  $a^{-2}$  converts coordinate distances into physical distances. Since  $\psi$  is related to the density perturbation by the Newtonian expression, it is called the peculiar gravitational potential. It is independent of  $t$  because, from Eq. (34),  $\delta\rho \propto a^2$ .

From Eq. (33) we see that the peculiar gravitational potential is related to the spatial curvature perturbation by

$$\psi = -\frac{3}{5}\mathcal{R}(\text{final}). \quad (49)$$

From Eqs. (46) and (48) the Fourier components of  $\mathbf{v}$ ,  $\psi$  and  $\delta$  are related by

$$\mathbf{v}_{\mathbf{k}} = i\frac{\mathbf{k}}{k} \left( \frac{aH}{k} \right) \delta_{\mathbf{k}} \quad (50)$$

$$\psi_{\mathbf{k}} = -\frac{3}{2} \left( \frac{aH}{k} \right)^2 \delta_{\mathbf{k}}. \quad (51)$$

## 4.2 Peculiar velocity in general relativity

The Newtonian picture has to be replaced by general relativity before matter domination, and even after matter domination on scales that are not far enough inside the horizon. To define peculiar velocity in this case one can proceed as follows [7, 36]. Start with the definition Eq. (17) of the velocity gradient, and first assume that the vorticity is negligible since angular momentum conservation ensures that it decays like  $(\rho + p)a^{-5}$ . Then one can show [36] that the velocity gradient perturbation is of the form

$$\delta u_{ij} = \partial_i v_j + \frac{1}{2} \dot{h}_{ij} \quad (52)$$

where  $\mathbf{v}$  is a globally defined peculiar velocity field as in the Newtonian picture and the extra term  $h_{ij}$  is transverse,  $\partial_i h_{ij} = 0$ , and traceless,  $\delta^{ij} h_{ij} = 0$ . The extra term represents the effect of gravitational waves, and at any epoch after matter domination the Newtonian picture is valid on scales well inside the horizon provided that it is negligible. It is certainly negligible at the present time in the region of the universe around us with radius a few hundred Mpc, where detailed galaxy surveys are performed, but it could be significant on bigger scales and at earlier times and so contribute to the cmb anisotropy [36]. Its presence does not however spoil Eqs. (45)–(51) because it is transverse and traceless. The conclusion, as advertized, is that the Newtonian equations of the last subsection apply to the peculiar velocity on all scales after matter domination. We shall use them in the next section to derive the Sachs-Wolfe effect.

One can avoid dropping the vorticity by proceeding as follows [36]. First, a transverse velocity  $\mathbf{v}^T$  may be *defined* by Eq. (44). The worldlines with velocity  $-\mathbf{v}^T$  relative to the

comoving worldlines have no vorticity, and comoving hypersurfaces are defined to be orthogonal to them (there are no hypersurfaces orthogonal to worldlines with nonzero vorticity). The velocity gradient  $\delta u_{ij}$  receives an extra contribution  $\frac{1}{2}(\partial_i v_j^T - \partial_j v_i^T) + \frac{1}{2}(\partial_i w_j^T + \partial_j w_i^T)$  where  $\mathbf{w}^T = \left[1 + 6 \left(1 + \frac{p}{\rho}\right) \left(\frac{aH}{k}\right)^2\right] \mathbf{v}^T$ . For scales well inside the horizon  $\mathbf{w}^T$  is negligible and the Newtonian result is recovered.

These results demonstrated, within the fluid flow approach, that the perturbations decouple into three modes. One is associated with gravitational waves, another with the vorticity of the velocity flow, and the third with the density perturbation that is our main concern.

## 5 The cmb anisotropy

The first detection by COBE of the intrinsic anisotropy of the cmb [1], announced in 1992, was surely the most important advance in observational cosmology for a long time. The COBE satellite explores large angular scales, corresponding to linear scales of order the particle horizon distance  $2H_0^{-1} \sim 10^4$  Mpc, and if the standard model is correct it provides a direct measurement of the primordial curvature perturbation. More recently other observations have detected smaller scale cmb anisotropy, which in due course will probe the finer details of the standard model and help to prove or disprove it [3, 4].

The anisotropy of the cmb is defined by giving the variation of its intensity with direction, at fixed wavelength. This variation is usually specified by giving the equivalent variation in the temperature of a blackbody distribution.<sup>10</sup> Denoting the direction of observation by a unit vector  $\mathbf{e}$ , the anisotropy may be expanded into multipoles

$$\frac{\Delta T(\mathbf{e})}{T} = \mathbf{w} \cdot \mathbf{e} + \sum_{l=2}^{\infty} \sum_{m=-l}^{+l} a_l^m Y_l^m(\mathbf{e}) . \quad (53)$$

The dipole term  $\mathbf{w} \cdot \mathbf{e}$  is well measured, and is the Doppler shift caused by our velocity  $\mathbf{w}$  relative to the rest frame of the cmb. Unless otherwise stated,  $\Delta T$  will denote only the intrinsic, non-dipole contribution from now on.

A feature in the sky of angular size  $\theta$  radians is dominated by multipoles of order  $l \sim 1/\theta$ . (This is analogous to the relation  $R \sim 1/k$  between the linear size of a feature and the wavenumbers that dominate its Fourier expansion.) Translating to degrees we have the following relation between  $l$  and the angular scale

$$\frac{\theta}{1^\circ} \simeq \frac{60}{l} . \quad (54)$$

The cmb originates at high redshift, and therefore on a comoving sphere whose present distance is close to the particle horizon at  $2H_0^{-1} = 6000h^{-1}$  Mpc. Its anisotropy also originates at comparable distances.<sup>11</sup> A feature with angular size  $\theta$  therefore corresponds to a linear scale

$$R \simeq 2H_0^{-1} \frac{\theta}{1 \text{ radian}} = 100h^{-1} \text{ Mpc} \frac{\theta}{1^\circ} . \quad (55)$$

Equivalently, the  $l$ th multipole of the cmb explores the scale

$$R \simeq 2H_0^{-1}/l = 6000h^{-1} \text{ Mpc}/l . \quad (56)$$

Present observations have resolution of order 1 degree, so they give information on scales  $100 \text{ Mpc} \lesssim k^{-1} \lesssim 10^4 \text{ Mpc}$ , and the lower limit will drop to around 10 Mpc when the angular resolution is improved. There is no significant anisotropy on smaller scales, because it is washed out by the finite thickness  $\simeq 7h^{-1} \text{ Mpc}$  of the last scattering surface.

<sup>10</sup>In principle the anisotropy depends on the wavelength but the dependence is not significant.

<sup>11</sup>Except for the Sunyaev-Zel'dovich effect, which is caused by galaxy clusters. It is significant only on arcminute scales, and will be ignored here.



## 5.1 The spectrum of the cmb anisotropy

The value of each multipole  $a_l^m$  depends on the observer's position. For a random choice of position there is some probability distribution, with zero mean, and some variance  $\langle |a_l^m|^2 \rangle_{\text{position}}$  which is the expected value of  $|a_l^m|^2$  for a randomly placed observer. The variance is independent of  $m$  because no direction in space is preferred. By analogy with Eq. (4), one may define the spectrum  $\mathcal{P}_T(l)$  of the cmb anisotropy by

$$\mathcal{P}_T(l) \equiv \frac{(2l+1)l}{4\pi} \langle |a_l^m|^2 \rangle_{\text{position}} . \quad (57)$$

The normalization is chosen so that the squared anisotropy averaged over both direction in the sky and position of the observer is

$$\left\langle \left\langle \left( \frac{\Delta T}{T} \right)^2 \right\rangle_{\text{sky}} \right\rangle_{\text{position}} = \sum_2^{\infty} \mathcal{P}_T(l)/l . \quad (58)$$

The cmb anisotropy is well described by linear theory, so the multipoles at a given position depend linearly on the Fourier components of the initial curvature perturbation  $\mathcal{R}$ . According to the standard model  $\mathcal{R}$  is Gaussian, which means essentially that the phases of its Fourier components are random. As a result the spectrum  $\mathcal{P}_T(l)$  depends linearly on the spectrum  $\mathcal{P}_{\mathcal{R}}(k)$  of the initial curvature perturbation, so there is a ‘photon transfer function’  $\mathcal{T}_l^2(k)$  such that  $\mathcal{P}_T(l) = \int_0^{\infty} \mathcal{T}_l^2(k) \mathcal{P}_{\mathcal{R}}(k) dk$ . Unless there is early re-ionisation, which we shall see is unlikely in viable versions of the standard model, the transfer function is determined by the amount and nature of the dark matter, through the same physical processes that we listed already. The only difference is that, except at the bottom end of the relevant range of scales  $10 \text{ Mpc} \lesssim k^{-1} \lesssim 10^4 \text{ Mpc}$ , it is not sensitive to the nature of the dark matter.

The spectrum  $\mathcal{P}_T(l)$  gives  $\langle |a_l^m|^2 \rangle_{\text{position}}$ , but we can only measure  $|a_l^m|^2$  at our position. The best guess is that it is equal to  $\langle |a_l^m|^2 \rangle_{\text{position}}$ , but one can also calculate the variance of this guess, which is called the *cosmic variance*. Because  $\mathcal{R}$  is Gaussian, the real and imaginary part of each multipole has an independent Gaussian distribution and as a result the cosmic variance of  $\sum_m |a_l^m|^2$  is only  $2/(2l+1)$  times its expected value. However, typical present observations suffer from an associated effect, *sample variance* [49], which arises because they cover only part of the sky preventing measurement of the full set of  $a_l^m$ . In essence, cosmic variance is the part of the sample variance which remains even once the full sky is observed, due to the recognition that complete determination of the spectrum would involve measurement on an ensemble of microwave skies rather than on the single one which is accessible.

A prediction for the spectrum is shown in Figure 1, and a qualitatively similar result is obtained in any viable version of the standard model.

## 5.2 The large scale cmb anisotropy (Sachs-Wolfe effect)

The simplest regime is the flat part of the spectrum corresponding to  $l \lesssim 30$  or angular scales bigger than a few degrees. If the standard model is correct the anisotropy in this regime probes directly the initial curvature perturbation, being caused essentially by the variations in the peculiar gravitational potential encountered by the cmb on its journey from the last scattering surface. This effect is called the Sachs-Wolfe effect, and we calculate it now for the case of critical density and zero cosmological constant following the approach of [7, 36]. The generalisation to the other cases [24, 25, 27] is completely straightforward, except for the lowest few multipoles where the effect of spatial curvature may become important [39].

It is convenient to separate the anisotropy into two parts,

$$\frac{\Delta T(\mathbf{e})}{T} = \left( \frac{\Delta T(\mathbf{e})}{T} \right)_{\text{initial}} + \left( \frac{\Delta T(\mathbf{e})}{T} \right)_{\text{jour}} . \quad (59)$$

The first term is the initial anisotropy, measured soon after decoupling by a set of comoving observers whose detectors point radially outwards and whose clocks are synchronized (this last condition just says that the anisotropy is defined on a comoving hypersurface). The second term is the additional anisotropy acquired on the journey towards us, including the Doppler shifts associated with the peculiar velocities of the initial comoving observers and of ourselves. (Our own peculiar velocity contributes only to the dipole but for the moment we are retaining all multipoles.)

On the scales in question the adiabatic initial condition Eq. (2) still holds at decoupling. Since  $\rho_r \propto T^4$  it follows that

$$\left( \frac{\Delta T(\mathbf{e})}{T} \right)_{\text{initial}} \simeq \frac{1}{3} \delta(\mathbf{x}_{\text{em}}, t_{\text{em}}) \quad (60)$$

where  $\delta$  is the matter density contrast. Here  $t_{\text{em}}$  denotes the time of emission (decoupling), and  $\mathbf{x}_{\text{em}} = 2H_0^{-1} \mathbf{e}$  denote the comoving coordinates of the point of emission.

To calculate the second term, consider a photon passing a succession of comoving observers [7, 36]. Its trajectory is  $a d\mathbf{x}/dt = -\mathbf{e}$  and between nearby observers its Doppler shift is

$$-\frac{d\lambda}{\lambda} = e_i e_j u_{ij} dr = -\frac{d\bar{a}}{\bar{a}} + e_i e_j \delta u_{ij} dr \quad (61)$$

where the first term is due to the average expansion, and the second is due to the relative peculiar velocity of the observers. Integrating this expression gives the redshift of radiation received by us, which was emitted from a distant comoving source. The unperturbed result is  $\lambda/\lambda_{\text{em}} = 1/a_{\text{em}}$ , and the first order perturbation gives

$$\left( \frac{\Delta T(\mathbf{e})}{T} \right)_{\text{jour}} = \int_0^{x_{\text{em}}} e_i e_j \delta u_{ij}(\mathbf{x}, t) a(t) dx . \quad (62)$$

The integration is along the photon trajectory

$$x(t) = \int_t^{t_0} \frac{dt}{a} = 3 \left( \frac{t_0}{a_0} - \frac{t}{a} \right) . \quad (63)$$

Neglecting any gravitational wave contribution,  $\delta u_{ij}$  is the gradient of the peculiar velocity, Eq. (43), which is given in terms of the peculiar gravitational potential by Eq. (46). Using Eq. (63) and integrating by parts one finds

$$\left( \frac{\Delta T(\mathbf{e})}{T} \right)_{\text{jour}} = \mathbf{e} \cdot [\mathbf{v}(0, t_0) - \mathbf{v}(\mathbf{x}_{\text{em}}, t_{\text{em}})] + \frac{1}{3} [\psi(\mathbf{x}_{\text{em}}) - \psi(0)] . \quad (64)$$

The term  $\mathbf{e} \cdot \mathbf{v}(0, t_0)$  is the Doppler shift corresponding to our peculiar velocity (in linear theory). It contributes only to the dipole and will be discounted from now on.<sup>12</sup> The term  $-\mathbf{e} \cdot \mathbf{v}(\mathbf{x}_{\text{em}}, t_{\text{em}})$  is the Doppler shift corresponding to the peculiar velocity of the initial observer. The remaining term therefore gives the anisotropy acquired on the journey towards us *if the initial and final anisotropies are defined by observers with zero peculiar velocity instead of by comoving observers*. It involves the gravitational potential  $\psi$  and can be thought of as being due to gravity, though we derived it by considering a sequence of Doppler shifts.

---

<sup>12</sup>By manipulating this expression it can be shown that in linear theory the dipole is precisely the Doppler shift due to our motion relative to the average motion of everything inside the sphere that emitted the cmb [36].

The term  $-\psi(0)/3$  can be dropped since it does not actually contribute to the anisotropy, leaving only the term  $\psi(\mathbf{x}_{\text{em}})/3$ .

Eq. (64) is exact in linear theory, but we are considering only scales that are far outside the horizon when the cmb is emitted. On these scales Eqs. (50) and (51) show that

$$|\psi_{\mathbf{k}}| \gg |\delta_{\mathbf{k}}| \gg |\mathbf{v}_{\mathbf{k}}|. \quad (65)$$

As a result we can drop the initial Doppler shift, and we can also drop the initial anisotropy Eq. (60). This gives the Sachs-Wolfe formula

$$\frac{\Delta T(\mathbf{e})}{T} = \frac{1}{3}\psi(\mathbf{x}_{\text{em}}). \quad (66)$$

It gives the large scale cmb anisotropy in terms of the peculiar gravitational potential, which is related to the curvature perturbation by  $\psi = -(3/5)\mathcal{R}$ . On the large scales that we are considering  $\mathcal{R}$  retains its initial value.

### The Sachs-Wolfe spectrum

Inserting the Fourier expansion of  $\mathcal{R}$  and projecting out the multipoles one finds the spectrum [51]

$$\mathcal{P}_T(l) = \frac{(2l+1)l}{25} \int_0^\infty \frac{dk}{k} j_l^2(2k/H_0) \mathcal{P}_{\mathcal{R}}(k) \quad (67)$$

where  $j_l$  is the spherical Bessel function and  $\mathcal{P}_{\mathcal{R}}$  is the spectrum of the initial curvature perturbation. For a power-law spectrum  $\mathcal{P}_{\mathcal{R}}(k) = (2k/H_0)^{n-1} \mathcal{P}_{\mathcal{R}}(H_0/2)$  one finds

$$\mathcal{P}_T(l) = \frac{1}{50} \frac{2l+1}{l+1} \left[ \frac{\sqrt{\pi}}{2} l(l+1) \frac{\Gamma((3-n)/2) \Gamma(l+(n-1)/2)}{\Gamma((4-n)/2) \Gamma(l+(5-n)/2)} \right] \mathcal{P}_{\mathcal{R}}(H_0/2). \quad (68)$$

For  $n = 1$  the square bracket is equal to 1.

### The COBE measurement of the cmb anisotropy

The COBE observations provide (after considerable analysis) estimates of the multipoles  $a_l^m$  in the range  $2 \leq l \lesssim 30$ . In this regime the Sachs-Wolfe formula applies leading to Eq. (68), and the multipoles are found to be consistent with this expression which is a highly non-trivial test of the standard model. From Eq. (68) one can estimate the spectral index  $n$  and the normalization  $\mathcal{P}_{\mathcal{R}}(H_0/2)$ . Instead of the latter one can specify any of the quantities  $\mathcal{P}_T(l)$ , which define the expected squared multipole. The usual quantity is the quadrupole, defined by

$$Q_{\text{rms-PS}}^2 \equiv 2T^2 \mathcal{P}_T(2). \quad (69)$$

From Eq. (58), the factor  $2T^2$  makes  $Q_{\text{rms-PS}}^2$  the quadrupole contribution to the expected mean square of  $\Delta T$ .

The latest analysis of the COBE data [2] finds that a good fit can be obtained with the spectral index anywhere in the range  $n \simeq 0.6$  to 1.4. At fixed  $n$  the normalization is however well determined, and in particular if  $n = 1$  then

$$Q_{\text{rms-PS}} = (19.9 \pm 1.6) \mu\text{K}, \quad (70)$$

which corresponds to  $\mathcal{P}_{\mathcal{R}}^{1/2} = 5.8 \times 10^{-5}$ .

The best fit value of  $Q_{\text{rms-PS}}$  depends significantly on  $n$ , but according to the analysis of [2] it becomes practically  $n$  independent if the 9th multipole is used instead. This fixes the normalization for all  $n$ .

### 5.3 The smaller scale cmb anisotropy

Once the normalisation has been determined, the smaller scale anisotropy can be calculated in a given version of the standard model. It is practically independent of the nature of the dark matter except on very small scales, and barring the possibility of early re-ionisation is therefore determined by  $n$ ,  $h$ , the matter density  $\Omega_m$  and any cosmological constant contribution  $\Omega_\Lambda$ . In the (critical density) MDM model that we are focussing on it therefore depends only on  $n$  and  $h$ .

As accurate measurements of the smaller scale cmb anisotropy are only beginning to be available firm conclusions are not yet possible, but this situation will change in the near future.

As shown in Figure 1, the predicted spectrum exhibits a series of peaks and troughs as one goes down in scale, which are usually called ‘Doppler peaks’. As an approximation, they may be regarded as coming from the standing-wave oscillation of the baryon-photon fluid density perturbation, between horizon entry and decoupling. The frequency of the oscillations goes up as the scale goes down, so the magnitude of the perturbation at decoupling is an oscillating function of scale. It causes both an initial anisotropy  $\Delta T_{\text{initial}}$  as seen by a comoving observer, and an additional anisotropy  $\Delta T_{\text{jour}}$  corresponding to the peculiar velocity of this observer.

Re-ionisation before  $z = 50$  or so would markedly reduce the height of the peaks, but as we discuss at the end of the next section such an early epoch is not predicted by viable versions of the standard model. The observations seem to be seeing the peaks at more or less the expected height [4] in weak confirmation of this prediction, but a quantitative analysis is not yet possible.

## 6 Theory and observation

Now we return to the density perturbation, and see what constraints are placed on it by observations of galaxies and clusters. We focus on  $\sigma(R, z)$ , the *rms* of the linearly evolved smoothed density contrast evaluated at an epoch corresponding to redshift  $z$ . As we noted earlier, apart from small corrections if a hot dark matter component is present it is proportional to  $(1+z)^{-1}$ , so it is enough to consider the present value

$$\sigma_0(R) \equiv (1+z)\sigma(R, z) . \quad (71)$$

Practically all of the data can be presented in terms of  $\sigma_0(R)$ , which makes it preferable to the more widely used power spectrum  $\mathcal{P}_0(k)$ . (A subscript 0 will always denote the present linearly evolved quantity.)

Some data points and theoretical curves are shown in Figures 2a and 2b, the latter being a close up of the former. The curves are all normalized to the COBE cmb anisotropy as described earlier. The full line corresponds to the canonical CDM model ( $n = 1$ ,  $h = 0.5$  and  $\Omega_\nu = 0$ ), using the CDM transfer function of reference [21] and taking  $\Omega_B = 0.05$  because of the nucleosynthesis relation  $\Omega_B h^2 = .013 \pm .002$ . The other lines show the effect of varying one at a time the parameters  $n$ ,  $h$  and  $\Omega_\nu$ . To vary  $\Omega_\nu$  at fixed  $h = 0.5$  we have used the parameterisation of the transfer function in reference [21], and to vary  $h$  at  $\Omega_\nu = 0$  we have used the approximation that the CDM transfer function depends only on  $kh^{-2}$ . (On small scales and in the presence of hot dark matter highly accurate transfer functions have yet to be published but these are probably adequate for the purpose at hand.)

Since the curves are rather similar and vary over a couple of orders of magnitude, the difference between them is best exhibited by normalising them all to a single model, which we take to be the canonical CDM model. The result is shown in Figure 3. In what follows the prediction of the canonical CDM will be denoted by  $\sigma_0^{\text{canon}}(R)$ .

To the extent that all of our COBE normalized curves come close to crossing at  $R \simeq 4000h^{-1}$  Mpc, the COBE data can be regarded as an observational point for  $\sigma_0(4000h^{-1}$  Mpc) and for convenience this is shown in the Figures. Its error bar is significant because it shows how much the theoretical curves can be moved up or down, but we emphasise that the normalization of each curve is determined directly from the cmb anisotropy as described above, not by fitting to this point.

The rest of the data points concern galaxies and clusters, as we now discuss.

## 6.1 The distribution of galaxies and galaxy clusters

The most extensive observations concern the distribution of galaxies and clusters, and provide information about  $\sigma_0(R)$  on scales  $10 \text{ Mpc} \lesssim R \lesssim 100 \text{ Mpc}$ . They also provide information about the correlation function  $\xi_0(R)$ . These quantities are related to the spectrum by

$$\sigma_0^2(R) = \int_0^\infty W^2(kR) \mathcal{P}_0(k) dk/k \quad (72)$$

$$\xi_0(R) = \int_0^\infty W(kR) \mathcal{P}_0(k) dk/k \quad (73)$$

where  $W$  is the ‘top hat’ window function, and  $\xi_0(R)$  is taken to be the volume averaged quantity [52]. A subscript 0 always denotes the present *linearly evolved* quantity.

The number density contrasts are known fairly well for IRAS galaxies, optical galaxies, radio galaxies and Abell galaxy clusters, out to a distance of several hundred Mpc. Within the rather large observational errors they are consistent with the biasing hypothesis,

$$\frac{\delta_I(\mathbf{x})}{b_I} \simeq \frac{\delta_O(\mathbf{x})}{b_O} \simeq \frac{\delta_R(\mathbf{x})}{b_R} \simeq \frac{\delta_A(\mathbf{x})}{b_A} \simeq \delta(\mathbf{x}) \quad (74)$$

where  $\delta_I$  etc. are the number density contrasts,  $\delta$  is the matter density contrast and  $b_I$  etc. are scale independent bias factors close to unity. The subscripts I, O, R and A refer to the four classes of objects mentioned. One can use the data to estimate the bias factors and hence the linearly evolved smoothed density contrast, on scales  $1 \text{ Mpc} \lesssim R \lesssim 100 \text{ Mpc}$ . (The determination of the overall normalization comes from nonlinear corrections, including redshift space to real space corrections.)

From a number of possibilities, we have chosen the analysis of Peacock and Dodds [52], which combines a variety of data sets. These authors find  $b_I : b_O : b_R : b_A = 1 : 1.3 : 1.9 : 4.5$  for the ratios of the bias factors, and  $b_I = 1.0 \pm 0.2$  for the overall normalization. Knowing the bias factors, the analyses of various groups gives estimates of  $\mathcal{P}_0(k)$ ,  $\sigma_0(R)$  and  $\xi_0(R)$ . Peacock and Dodds convert the last two into estimates of  $\mathcal{P}_0(k)$  using the prescription<sup>13</sup>

$$\sigma_0(R) = \mathcal{P}_0^{1/2}(k_R) \quad (75)$$

$$\xi_0(R) = \mathcal{P}_0^{1/2}(\sqrt{2}k_R) \quad (76)$$

where

$$k_R = \left[ \frac{1}{2} \Gamma \left( \frac{m+3}{2} \right) \right]^{1/(m+3)} \frac{\sqrt{5}}{R} \quad (77)$$

and  $m \equiv (k/\mathcal{P}_0)(d\mathcal{P}_0/dk)$  is the effective spectral index evaluated in the canonical CDM model. These formulae are obtained by taking  $m$  constant, and using the approximation

$$W(kR) = \exp(-k^2 R^2/10) \quad (78)$$

<sup>13</sup>Although this prescription relating  $\sigma_0(R)$  to  $\mathcal{P}_0(k)$  is adequate on scales  $\gtrsim 10 \text{ Mpc}$ , it is too dependent on the shape of the transfer function to be useful on smaller scales. As we shall see, data on such scales directly constrain  $\sigma_0(R)$ , which is our main reason for focusing on that quantity rather than on  $\mathcal{P}_0(k)$  or  $\xi_0(R)$ .

which is exact for  $kR \ll 1$ . We have used this same prescription to convert the Peacock-Dodds estimates of  $\mathcal{P}_0(k)$  into estimates of  $\sigma_0(R)$ , and the results are shown as open squares in Figures 2 and 3. For clarity a point with particularly large error bars, which lies between the two largest scale points shown, has been omitted.

Most of the uncertainty in this determination of  $\sigma_0(R)$  comes from the uncertainty in the *overall* normalization, defined by  $b_I = 1.0 \pm 0.2$ . If all of the raw data used in the determination referred directly to real space number densities,  $\sigma_0$  would be proportional to  $b_I^{-1}$  times the raw data, so its uncertainty would be the same as for  $b_I$ , namely 20%. This is the case for the results of the APM survey, which is a very important part of the total input, but as Peacock and Dodds note the rest of the input is based on redshift space number densities. In the linear regime the determination of  $\sigma_0$  from raw data on a given type of object in redshift space is proportional to  $[b_I(1 + .66rb_I^{-1} + .20r^2b_I^{-2})]^{-1}$ , where  $r \equiv b_I/b_N$ . For clusters  $r = 4.5$  so that  $\sigma_0$  is more or less proportional to  $b_I^{-1}$  as for real space data, but for the three types of galaxy  $r = 1.0, 1.3$  and  $1.9$  and  $\sigma_0$  is roughly proportional to  $b_I^{-5/7}$ . Thus, galaxy data in redshift space in the linear regime gives an uncertainty in  $\sigma_0(R)$  of order  $(5/7) \times 20\% = 14\%$ . To be on the safe side we work with 20%, noting that at least in the linear regime the true uncertainty is between 14% and 20%.

With  $b_I$  fixed, one obtains comparatively small error bars, which are indicated by the inner error bars in Figures 2 and 3. As an aid to clarity, in Figure 3 we have connected the ends of these error bars by straight lines to form a box. Within the box the points can be moved up and down more or less independently, and the whole box can be moved up or down by 20% as indicated by the outer error bars. Both these types of error are the result of statistical fitting, and presumably correspond to something like a 1- $\sigma$  level (with the caveat that the outside error bars might be reduced somewhat as discussed above).

## 6.2 The bulk flow

The smoothed peculiar velocity field  $\mathbf{v}(R, \mathbf{x})$  is called the bulk flow, and on scales above 10 Mpc it is described by linear theory. It is therefore the gradient of a potential, and can in principle be constructed from its radial component [53]. The latter is directly observed from the redshift-to-distance ratios of galaxies, in the region around us with radius a few hundred Mpc.

Unlike the density contrast, the bulk flow can reasonably be assumed to be the same as that of the underlying matter at least on large scales, which in principle allows one to dispense with the biasing hypothesis. In practice one still needs the hypothesis at present if the data are to yield really powerful results, which are obtained by comparing the density field obtained from the bulk flow with that obtained from galaxy surveys. A recent study [54] concludes that at 95% confidence level  $b_I = 0.7^{+0.6}_{-0.2}$ .

Although the bulk flow alone does not yet give very powerful results, it is not completely useless. The standard way of utilising it is to look at the magnitude  $v(R, 0)$  of the bulk flow at our position, or in other words at the average peculiar velocity in the sphere around us of radius  $R$ .<sup>14</sup> From Eq. (50) the expected value of  $v^2(R, 0)$  is

$$\sigma_v^2(R) = \int_0^\infty \left(\frac{H_0}{k}\right)^2 W^2(kR) \mathcal{P}_0(k) \frac{dk}{k} . \quad (79)$$

Evaluated at random positions, each of the three components of the bulk flow has a Gaussian distribution with one third of this variance, so at 1- $\sigma$  level one expects that

$$\frac{v(R, 0)}{\sigma_v(R)} = 1^{+32\%}_{-52\%} . \quad (80)$$

---

<sup>14</sup>There are several variants of this procedure, which correspond to using window functions different from a top hat. In what follows we are actually using the variant described in [7].

Observational values of  $v(R, 0)$  are available only for  $R$  of order tens of Mpc, and nothing is lost by considering just  $R = 40h^{-1}$  Mpc. The canonical CDM model gives [6]  $\sigma_v^{\text{canon}}(40h^{-1} \text{ Mpc}) = 460 \text{ km sec}^{-1}$ , and the observational value is  $v_{\text{obs}}(40h^{-1} \text{ Mpc}, 0) = (400 \pm 15\%) \text{ km sec}^{-1}$ . This error is swamped by the ‘cosmic variance’ given by Eq. (80), and ignoring it we deduce that

$$\frac{\sigma_v(40h^{-1} \text{ Mpc})}{\sigma_v^{\text{canon}}(40h^{-1} \text{ Mpc})} = (400/460)^{+92\%}_{-24\%} \quad (81)$$

where the error is just the inverse of that in Eq. (80).

Compared with Eq. (40) for  $\sigma_0(R)$ , Eq. (79) contains an extra factor  $(H_0/k)^2$  and as a result  $\sigma_v(40h^{-1} \text{ Mpc})$  probes about the same scale  $k^{-1}$  as  $\sigma_0(90h^{-1} \text{ Mpc})$ . We conclude that the above ratio applies to this latter quantity, which leads to the estimate  $\sigma_0(90h^{-1} \text{ Mpc}) = .044^{+.040}_{-.010}$  shown in Figures 2 and 3.

### 6.3 Abundances

The rest of our observational points invoke the well known Press-Schechter estimate [7] for the contribution to the mass density of objects with mass  $> M$ ,

$$\Omega(> M(R), z) = \text{erfc} \left( \frac{\delta_c}{\sqrt{2}\sigma(R, z)} \right) \quad (82)$$

where  $\delta_c = 1.7$  and  $M(R)$  is the mass in a comoving sphere of present radius  $R$ , given by Eq. (38). Except for a more or less unmotivated factor 2 the right hand side is just the fraction of space occupied by regions of space where the linearly evolved smoothed density contrast exceeds  $\delta_c$ , so roughly speaking the Press-Schechter estimate simply states that such regions are occupied by objects with mass bigger than  $M$ . It is only supposed to be valid as long as these regions are rare, corresponding to  $\Omega(> M, z) \ll 1$ . In this regime  $\sigma(R, z) \ll 1$ , and linear evolution should be valid.

The value  $\delta_c = 1.7$  is motivated by a spherical collapse model, and N-body simulations suggest that the Press-Schechter estimate is roughly correct, if  $\delta_c$  is within ten percent or so of this value [55, 56].<sup>15</sup>

In some cases the observational result is given directly as a value of  $\Omega(> M)$ , but in others it is given as a number density  $n(> M)$ . In the latter case we obtain an observational value of  $\Omega(> M)$  by setting  $\rho(> M) \simeq Mn(> M)$  which gives

$$\Omega(> M) = 3.60h \frac{M}{10^{12} M_\odot} \frac{n(> M)}{(h^{-3} \text{ Mpc})^{-3}}. \quad (83)$$

This overestimates  $\Omega(> M)$  and therefore  $\sigma(R)$ , but in practice the mass of the observed objects is too uncertain for this to matter very much.

An alternative is to obtain a theoretical prediction for  $n(> M)$  from the Press-Schechter formula. This more complicated procedure has been used by many authors (starting with the originators of the formula), but because of the uncertainty in the mass there is generally no significant advantage in using it. Some relevant comparisons are cited below.

### Galaxy clusters

We first look at galaxy clusters. The most useful are those of ‘richness class  $R > 1$ ’, which are the heaviest bound objects. Their present number density is known to be about  $n \simeq 8 \times 10^{-6} h^3 \text{ Mpc}^{-3}$  and their mass  $M_{\text{clus}}$  is roughly  $10^{15} M_\odot$ . The uncertainty

<sup>15</sup>Somewhat smaller values are found with a Gaussian instead of a top-hat smoothing but we are using the latter.

in the mass is the limiting factor in the present context. Let us first set it equal to  $1.2 \times 10^{15} M_{\odot}$  which corresponds (with  $h = .5$ ) to  $R = 8h^{-1}$  Mpc. Then Eqs. (82) and (83) give  $\sigma_0(8h^{-1} \text{ Mpc}) = 0.71$ . Lowering  $M_{\text{clus}}$  by a factor 2 gives  $\sigma_0(6.4h^{-1} \text{ Mpc}) = 0.64$ , and raising it by a factor 2 gives  $\sigma_0(10.1h^{-1} \text{ Mpc}) = 0.80$ . These three values are indicated in Figures 2a and 2b by a central point and a slanting ‘error bar’. Because  $\sigma_0^{\text{canon}}(R)$  is strongly varying the corresponding uncertainty in  $\sigma_0(R)/\sigma_0^{\text{canon}}(R)$  is quite big, namely +43% and -27%. Compared with this uncertainty the range of scales  $R = 6.4h^{-1}$  to  $10h^{-1}$  Mpc is negligible (essentially because all the spectra we consider have broadly similar shapes), and we have ignored the horizontal error when plotting Figure 3.

Our conclusion is that if  $R > 1$  clusters have mass  $M = 10^{15} M_{\odot}$  within a factor 2, the normalization of the standard model of structure formation is  $\sigma_0(8h^{-1} \text{ Mpc}) = 0.71^{+43\%}_{-27\%} = 0.5$  to 1.0. Estimates of the normalization in a similar spirit have been made by previous authors. A recent and typical one [57] quoted a normalization  $\sigma(8h^{-1} \text{ Mpc}) \simeq 0.57$ , but did not quantify the effect of the uncertainty in the cluster mass.

One can in principle apply the same technique to clusters observed at high redshift. A treatment of  $z = 0.3$  observations [58] which estimates the effect of both mass and number density uncertainty gives  $\sigma_0(8h^{-1} \text{ Mpc}) = 0.6$  to 0.9, in excellent agreement with our result. High redshift observations of clusters are likely to become increasingly important in the future.

## Quasars and damped Lyman alpha systems

The same technique can be applied to quasars and damped Lyman alpha systems observed at redshift 3 to 4. Their masses are thought to be roughly in the range  $10^{11}$  to  $10^{13} M_{\odot}$  and on these scales the Press-Schechter formula gives  $\Omega(> M, z) \ll 1$ , indicating that linear theory applies and the formula is valid. An observational value of the quasar fraction  $\Omega_{\text{quas}}(> M, z)$  or the damped Lyman alpha system fraction  $\Omega_{\text{lyalph}}(> M, z)$  therefore provides a lower limit on  $\Omega(> M, z)$ , and on  $\sigma_0(R)$ .<sup>16</sup>  $\Omega_{\text{quas}}(> M, z)$  can be estimated from observation out to  $z = 4$ , in the range roughly  $10^{11} M_{\odot}$  to  $10^{13} M_{\odot}$ , if one knows the baryon fraction that forms the central black hole of each quasar and some other astrophysical parameters. Haehnelt [59] takes 1% as a reasonable estimate of the fraction, and using reasonable values for the other parameters estimates that  $\sigma_0(M = 10^{13} M_{\odot}) > 1.30$ . Multiplying the baryon fraction by 10 (or making equivalent changes in the other parameters) reduces this to  $\sigma_0(M = 10^{13} M_{\odot}) > 1.25$ , which is shown in figures 2 and 3.<sup>17</sup> One sees that this is not very constraining. He also gives bounds for lower masses, down to  $M = 10^{11} M_{\odot}$  but they are even less restrictive. (A more recent estimate [56] using newer data suggests that a better constraint might now be possible, but a detailed analysis has not yet been done.)

More restrictive bounds [18] come from damped Lyman alpha systems. The data indicate [60] that the *baryons* in such systems account for  $\Omega_{\text{lyalph}}(> M, (3.0 \pm 0.5)) = (0.0029 \pm 0.0006)h^{-1}$ . Let us take the lower bound of the redshift and  $\Omega$  to obtain a lower bound on  $\sigma_0$  (we set  $h = 0.5$  as its uncertainty is not significant). Then if the baryon fraction is the same as the universal fraction  $\simeq 0.05$ , we deduce that  $\Omega(> M_{\text{lyalph}}, 3.0) > 0.092$  and  $\sigma_0(R) > 4.0$  where  $R$  is the scale corresponding to  $M_{\text{lyalph}}$ . To obtain a reasonably firm bound let us divide this result by 10 (corresponding, say, to assuming that the baryons are concentrated by a factor 10). This leads to  $\sigma_0(R) > 2.6$ .

The corresponding bound on  $\sigma_0(R)/\sigma_0^{\text{canon}}(R)$  depends on  $M_{\text{lyalph}}$ , but the dependence is much weaker than it is for clusters because  $\sigma_0^{\text{canon}}(R)$  is varying more slowly with scale. The bounds displayed in Figures 2 and 3 are for  $M_{\text{lyalph}} = 3 \times 10^{11} M_{\odot}$ , corresponding to

<sup>16</sup>For the purpose of this article ‘quasar mass’ is taken to mean the total mass of the object concerned, though the term ‘quasar’ usually refers only to the region around the central black hole.

<sup>17</sup>The weak dependence on the baryon fraction arises because  $\Omega_{\text{quas}}(> 10^{13} M_{\odot}, 3)$  is tiny, so that the prediction for it is an exponentially decreasing function of  $\sigma_0$ .



$R = 0.50h^{-1}$  Mpc. Going down to  $10^{10}M_{\odot}$  only lowers  $\sigma_0(R)/\sigma_0^{\text{canon}}(R)$  by 17% and going up to  $10^{12}M_{\odot}$  only raises it by 7%, the corresponding scales being  $R = 0.16h^{-1}$  Mpc and  $R = 0.76h^{-1}$  Mpc. None of these changes is very significant.

Our Lyman alpha constraint corresponds to a constraint  $\Omega_{\nu} \lesssim 0.2$  (with  $n = 1$  and  $h = 0.5$ ). This is the same as that obtained in most (but not all) previous investigations [18, 22], which have used both the Press-Schechter formula and numerical simulations.

## 6.4 Other observations

The observations that we have considered are the only ones that can be related more or less directly to the linearly evolved density contrast. Other observations can in principle yield information after comparing them with numerical simulations, but at present the uncertainties seem to be too big to make the additional information useful. For example, the pairwise galaxy velocity dispersion has been widely compared with numerical simulations and the received wisdom is that with the COBE normalization the prediction is significantly too high (eg. [16]). But others [61] argue that when the observational [62] and theoretical uncertainties are taken into account the discrepancy ceases to be significant.

## 6.5 Summary of the data on $\sigma_0(R)$

The most accurate data point is the one coming from the COBE measured CMB anisotropy, which is indicated schematically by the point at  $4000h^{-1}$  Mpc. That is, of course, the reason why we have used it to normalize the theoretical curves.

The other points all come from galaxy and cluster observations, and for clarity we will focus on  $\sigma_0(R)/\sigma_0^{\text{canon}}(R)$  as displayed in Figure 3. Most of the points come from the number density contrasts of various types of object, as analysed by Peacock and Dodds, which are represented by open squares. Within the box each point can be moved down more or less independently, and the whole box can be moved up about 20% as indicated by the outer error bars. (Both of these uncertainties are more or less at the  $1\text{-}\sigma$  level, though the 20% estimate may be somewhat high according to this criterion.) The most striking feature of these points is the slope of  $\sigma_0(R)/\sigma_0^{\text{canon}}(R)$ . In the regime  $10h^{-1} \text{ Mpc} \lesssim R \lesssim 40h^{-1} \text{ Mpc}$  it is positive and rather well defined. This positive slope is seen in both the galaxy correlation function (in many different surveys) and in the cluster correlation function [63], and in addition it seems to be needed to make N-body simulations give the correct *normalization* for the cluster correlation function [63]. In other words, it is rather firmly established. On smaller scales the slope is flatter, and there is also a hint of flattening on larger scales.

What about the overall normalization of these points? The 20% uncertainty indicated by the outer error bars corresponds to at least a  $1\text{-}\sigma$  confidence level in the context of the Peacock and Dodds analysis, and both the upper and lower limits do have confirmation from other types of data. The *lower* limit corresponds to the *upper* limit of Peacock and Dodds' bias factor determination  $b_I = 1.0 \pm 0.2$ , which is confirmed by the result [54]  $b_I = 0.7^{+0.6}_{-0.2}$  at 95% confidence level, that comes from combining the galaxy distribution with observations of the bulk flow. The lower limit of the normalization is also confirmed, less strongly, by the bulk flow data on the scale  $R \sim 90h^{-1}$  Mpc. Coming to the upper limit on the normalisation of  $\sigma_0(R)/\sigma_0^{\text{canon}}(R)$ , one sees from Figure 3 that it is confirmed on the scale  $R \sim 10h^{-1}$  Mpc by the abundance of present day rich galaxy clusters if one assumes that their mass is less than  $2 \times 10^{15}M_{\odot}$ . A mass in this range is certainly commonly assigned to these rich clusters, and indeed a lower (median) mass corresponding to a lower upper limit on  $\sigma_0(8h^{-1} \text{ Mpc})$  has been advocated [57]. Essentially the same upper limit is confirmed by the abundance of  $z = 0.3$  clusters as cited earlier.

All this is on scales 10 to 100 Mpc. A crucial lower bound on the normalization on some scale in the decade 0.1 to 1 Mpc is provided by the abundance of damped Lyman

alpha systems at  $z = 3$ . In contrast with the case of galaxy clusters, the uncertainty about the precise scale matters little here, because the theoretical curves that one is trying to normalise are almost flat. It is noteworthy that this lower limit is already presaged by the flattening of the slope seen in the regime  $4h^{-1} \lesssim R \lesssim 10h^{-1}$  Mpc.

## 6.6 Theory versus observation

The first conclusion to draw from the comparison of theory and observation is that the canonical CDM model does not work. As discussed in the Introduction three fixes are possible. One is an ‘old age’ model where  $h$  is very low. Another is a ‘tilted’ model where  $n$  is significantly below 1. The third is to alter the nature or the amount of the cold dark matter, and the specific option that we have explored is to introduce some fraction  $\Omega_\nu$  of hot (neutrino) dark matter while keeping the total matter density fixed at the critical value.

Figure 3 shows that any of these three options is capable of giving an acceptable fit to the data, the required parameters being  $h \simeq 0.3$  for the first option,  $n \simeq 0.7$  for the second and  $\Omega_\nu \simeq 0.15$  to  $0.20$  for the third. In each case more extreme values (smaller  $h$  or  $n$  or bigger  $\Omega_\nu$ ) are ruled out by the observational lower bounds on the normalization of  $\sigma_0(R)$ , especially the one coming from damped Lyman alpha systems (recall that the one we show is supposed to be rather firm). Less extreme values are ruled out by the *slope* of  $\sigma_0(R)$  (and in the case of the MDM option also by the upper limit on its magnitude) in the tens of Mpc range.

What can we say about the relative merits of the three options? As noted earlier, in most (particle physics motivated) models of inflation that implement the tilted option there is also a gravitational wave contribution whose effect is to lower the COBE normalization by  $[1 + 6(1 - n)]^{-1/2}$ . For  $n = 0.7$  this factor is 0.6 which leads to a gross contradiction with the data. As a result the tilted option is not viable in most models of inflation, the only known exception being ‘natural’ inflation [64, 15]. The old age option is also problematical because of the low Hubble constant that it invokes, which leaves the MDM option as perhaps the most attractive of the three.

So far everything is quite simple, but we have only allowed the three parameters to depart from their canonical values  $h = 0.5$ ,  $n = 1.0$  and  $\Omega_\nu = 0$  one at a time. We would like to understand the entire parameter space. In particular we would like to know the observational constraint on  $n$  because, especially in conjunction with gravitational waves, it is such a useful discriminator between models of inflation.

One way forward [20, 21, 23] is to perform a least squares fit, though this is somewhat problematical because the damped Lyman alpha systems provide only a bound, and because it is not clear that all of the points coming from galaxy and cluster correlations are independent. An alternative [8] is to simply demand that the curves go through selected points, and have slopes within the limits indicated by the box. These options will be explored in a future paper [65], but for the moment we point out that rather definite conclusions [10] about  $n$  can already be drawn just from Figure 3.

To start with, note that in the regime where we have data, a change in  $h$  is roughly equivalent to a change in  $n$  according to the formula  $\Delta h = -\Delta n$ . This can be seen by comparing the  $h = 0.3$  and  $n = 0.7$  curves in Figure 3, especially for the crucial slope. Accepting this and taking  $0.4 \lesssim h \lesssim 0.6$ , one concludes that with critical density CDM requires  $n \simeq 0.6$  to  $0.8$ .

The crucial point now is that *all* known modifications of the pure, critical density CDM hypothesis *reduce* the COBE normalized  $\sigma_0(R)$ , because CDM maximises the growth of the density perturbation on small scales. Thus the *lower* limit  $n \gtrsim 0.6$  will hold in *any* version of the standard model. If gravitational waves reduce the COBE normalization by a factor  $[1 + 6(1 - n)]^{-1/2}$ , it is not difficult to verify that the lower limit is tightened to  $n \gtrsim 0.8$ .

What about the upper limit? On scales  $R \lesssim 40h^{-1}$  Mpc or so one can clearly cancel an

Table 1: Estimates of the re-ionisation epoch  $1 + z_{\text{ion}}$ 

$h$	$n$	$\Omega_\nu$	low estimate	best guess	high estimate
0.5	1.0	0	31	53	78
0.5	0.7	0	9.0	15	23
0.5	1.0	0.3	3.9	7.6	12
0.3	1.0	0	6.1	10	15

increase in  $n$  by increasing the fraction of hot dark matter, but at least within the critical density MDM option this is impossible on larger scales. Assuming that no cancellation is possible, the upper bound  $n \lesssim 1$  indicated by the four largest scale galaxy/cluster correlation points shown in Figure 3 can only be evaded by decreasing  $h$ , which (assuming that  $h > 0.4$ ) allows it to go up to 1.1. A similar upper bound comes from the fact that too much hot dark matter gives a strong violation of the damped Lyman alpha system bound, which cannot be cancelled by increasing  $n$  because of the data in the tens of Mpc range. The conclusion is that, at least within the MDM option, there is an upper bound  $n \lesssim 1.1$  or so.

### The epoch of re-ionisation

We end with a brief discussion of estimates of the re-ionisation epoch  $z_{\text{ion}}$ , which are as yet rather preliminary.

Observationally, a lower limit of order  $z_{\text{ion}} \gtrsim 5$  is well established [66], and a useful upper limit or even a value may soon become available from the small scale cmb anisotropy [67].

To estimate the re-ionisation epoch theoretically, one starts by estimating the fraction  $f$  of the mass that has to collapse into galaxies before re-ionisation occurs. At present the astrophysics is not understood sufficiently well to allow a firm calculation of  $f$ . Tegmark, Silk and Blanchard [67] give a ‘best guess’  $f \simeq 4 \times 10^{-3}$  with upper and lower limits  $2 \times 10^{-5} \lesssim f \lesssim 0.4$ .

If  $M_{\text{min}}$  is the mass of the smallest galaxies existing at the epoch of re-ionisation, then  $\Omega(> M_{\text{min}}, z_{\text{ion}}) = f$ , and the Press-Schechter estimate gives  $z_{\text{ion}}$  in terms of  $\sigma_0(R_{\text{min}})$  and  $f$ . We take as a best guess  $M_{\text{min}} = 10^6 M_\odot$  (the same as the minimum mass of present day galaxies), while noting that theoretical estimates vary from  $10^5 M_\odot$  to  $10^8 M_\odot$  [67]. The corresponding scale is  $R_{\text{min}} \simeq 0.01$  Mpc, two orders of magnitude smaller than the scales that we have considered so far.

If  $f \sim 1$  re-ionisation will not occur until the epoch of non-linearity given by  $(1 + z_{\text{nl}}) = \sigma_0(R_{\text{min}})$ , but if  $f$  is smaller it will be earlier. Tegmark *et al*’s ‘best guess’ gives  $(1 + z_{\text{ion}}) \simeq 1.7\sigma_0(R_{\text{min}})$  whereas the lower end of their allowed range gives  $(1 + z_{\text{ion}}) \simeq 2.5\sigma_0(R_{\text{min}})$ . The range is not too big, because if  $f$  is small it is an exponentially decreasing function of  $\sigma_0(R_{\text{min}})$ .

Tegmark *et al* go on to calculate  $\sigma_0(0.01 \text{ Mpc})/\sigma_0(8h^{-1})$  in various models. Inserting the values of  $\sigma_0(8h^{-1})$  following from our COBE normalization we deduce the results in Table 1. The low estimate of  $z_{\text{ion}}$  takes it to be the epoch of non-linearity, corresponding to  $f \sim 1$ , the other two correspond to the values of  $f$  mentioned earlier. In the third row we have allowed for the slower growth of the MDM density contrast in the same way as Tegmark *et al* (following [19]), and in the last row we have used the scaling  $\sigma_0(R) \propto h^2$  that is appropriate in the small scale regime where  $\sigma_0(R)$  is practically independent of  $R$ . Note that the third row is for 30% hot dark matter, whereas at most 20% or so is viable. Thus a viable MDM option will give earlier re-ionisation than the other two options.

Although the uncertainties are big, it seems clear that viable versions of the standard

model do not lead to very early re-ionisation. On the other hand a re-ionisation redshift in the low tens, which might be big enough to be affect the cmb anisotropy, is not out of the question.

## 7 Conclusion

To conclude, we have provided a review of the machinery required in order to translate an initial spectrum of density irregularities into a form amenable for fairly direct comparison with a range of observations. Although some of our discussion has wider application, we have focussed on the standard model of structure formation, which invokes an initially Gaussian, adiabatic, scale invariant density perturbation and more or less cold dark matter. It is at present the one under the most active consideration, because it is the simplest, and also because the required initial density perturbation might be generated as a vacuum fluctuation during inflation.

We have gone on to compare the standard model with presently available data, indicating the relative merits of altering the Hubble constant, tilting the primordial spectrum and incorporating a component of hot dark matter. All of these have been envisaged as ways to remedy the shortcomings of the canonical cold dark matter model.

Our treatment of the observations breaks new ground in that it combines for the first time several relatively new observational constraints. For the normalization of the spectrum from the COBE measurement of the cmb anisotropy, we use the most recent determination that has been provided by Górski and collaborators. It is significantly higher than earlier determinations. We assess carefully the uncertainty in the determination of the normalization on the scale  $8h^{-1}$  Mpc that comes from the cluster abundance, quantifying its dependence on the assumed cluster mass. We establish a simple and rather firm constraint from damped Lyman alpha systems at high redshift, which provide a powerful lower limit on some scale of order 0.1 to 1 Mpc. Finally, on the basis of an earlier study [67] we make a preliminary estimate of the epoch of re-ionisation in the various models.

Our technique for comparing theory with observation is to look not at the power spectrum, but at the dispersion of the density contrast smoothed on a scale  $R$ . This has the great advantage of being closer to what is actually measured, while still being simple to calculate for a given theoretical model. Particularly when one normalises theories and observations to a benchmark model as in Figure 3, the comparison between theory and observation can be very clearly illustrated.

While present observations are strong enough to exclude the standard CDM model, they are not of sufficient quality to select amongst different ways of generalising it, particularly when one realises that if tilt or a hot component are to be varied then one must certainly also allow  $h$  to vary in combination. It is clear from Figure 3 that these options are most cleanly probed by observations on scales from a few tens to a few hundreds of megaparsecs, and we should look forward to this region of the spectrum being probed by both small scale microwave anisotropy experiments and by larger scale galaxy correlation (and peculiar velocity) measurements.

## Acknowledgements

ARL is supported by the Royal Society, and acknowledges use of the Starlink computer system at the University of Sussex. It is a pleasure to thank Bob Schaefer, Qaisar Shafi and Pedro Viana for many discussions.

## References

- [1] G. F. Smoot *et al.*, *Astrophys. J. Lett.* **396** (1992) L1.
- [2] K. Górski *et al.*, “On determining the spectrum of primordial inhomogeneity from the COBE DMR sky maps: II. Results of two year data analysis”, COBE preprint (1994).
- [3] M. White, D. Scott and J. Silk, to appear in *Ann. Rev. Astron. Astroph.* (1994).
- [4] D. Scott and M. White, “The existence of baryons at  $z = 1000$ ”, Berkeley preprint (1994).
- [5] A. R. Liddle, D. H. Lyth and W. J. Sutherland, *Phys. Lett.* **B279** (1992) 244.
- [6] A. R. Liddle and D. H. Lyth, *Phys. Lett.* **B291** (1992) 391.
- [7] A. R. Liddle and D. H. Lyth, *Phys. Rep.* **231** (1993) 1.
- [8] A. R. Liddle and D. H. Lyth, *Mon. Not. Roy. astr. Soc.* **265** (1993) 379.
- [9] A. R. Liddle and D. H. Lyth, *Annals of the New York Academy of Sciences* **688** (1993) 653.
- [10] D. H. Lyth and A. R. Liddle, to appear in *Astrophys. Lett. & Commun.* (1994) (proceedings of the Capri workshop); D. H. Lyth, in *Proceedings of the 1993 Trieste Summer School* (World Scientific Press, Singapore, 1994).
- [11] R. L. Davis, H. M. Hodges, G. F. Smoot, P. J. Steinhardt and M. S. Turner, *Phys. Rev. Lett.* **69** (1992) 1856.
- [12] D. S. Salopek, *Phys. Rev. Lett.* **69** (1992) 3602.
- [13] J. G. Bartlett *et al.*, “The case for a Hubble constant of  $30 \text{ km sec}^{-1} \text{ Mpc}^{-1}$ ”, preprint (1994).
- [14] N. Vittorio, S. Matarrese and S. Lucchin, *Astrophys. J.* **328** (1988) 69; J. P. Ostriker and Y. Suto, *Astrophys. J.* **348** (1990) 378; J. R. Bond, in *Highlights in Astronomy* vol. 9 Proceedings of the IAU Joint discussion, Ed. J. Bergeron (Buenos Aires Gen Ass) (1992); R. Cen, N. Y. Gnedin, L. A. Kofman and J. P. Ostriker, *Astrophys. J.* **399** (1992) L11; D. S. Salopek, *Phys. Rev.* **45** (1992) 1139.
- [15] F. C. Adams, J. R. Bond, K. Freese, J. A. Frieman and A. V. Olinto, *Phys. Rev.* **D47** (1993) 426.
- [16] J. M. Gelb, B.-A. Gradwohl and J. A. Frieman, *Astrophys. J. Lett.* **403** (1993) L5.
- [17] S. A. Bonometto and R. Valdarnini, *Phys. Lett.* **103A** (1984) 369; Q. Shafi and F. W. Stecker, *Phys. Rev. Lett.* **53** (1984) 1292; L. Z. Fang, S. X. Li and S. P. Xiang, *Astron. Astrophys.* **140**, (1984) 77; R. Valdarnini and S. A. Bonometto, *Astron. Astrophys.* **146** (1985) 235; S. Achilli, F. Occhionero & R. Scaramella, *Astrophys. J.* **299** (1985) 577; S. Ikeuchi, C. Norman & Y. Zhan, *Astrophys. J.* **324**, (1988) 33; R. K. Schaefer, Q. Shafi and F. Stecker, *Astrophys. J.* **347** (1989) 575; J. Holtzman, *Astrophys. J. Supp.* **71** (1989) 1; E. L. Wright *et al.*, *Astrophys. J. Lett.* **396** (1992) L13; R. K. Schaefer and Q. Shafi, *Nature* **359** (1992) 199; A. N. Taylor and M. Rowan-Robinson, *Nature*, **359** (1992) 396; T. van Dalen and R. K. Schaefer, *Astrophys. J.* **398** (1992) 33; J. Holtzman and J. Primack, *Astrophys. J.* **405** (1993) 428; A. Klypin, J. Holtzman, J. R. Primack, and E. Regös, *Astrophys. J.* **416** (1993) 1; Y. P. Jing, H. J. Mo, G. Börner & L. Z. Fang, to appear, *Astron. and Astrophys.* (1994); S. Ghigna *et al.*, “Void analysis as a test for dark matter models”, preprint (1994); G. Yepes *et al.*, “The angular correlation function of galaxies in cdm and chdm models”, preprint (1994).

- [18] K. Subramanian and T. Padmanabhan, , “Constraints on the models for structure formation from the abundance of damped lyman alpha systems”, IUCAA preprint (1994); H. J. Mo and J. Miralda-Escude, “Damped lyman alpha systems and galaxy formation”, Princeton preprint (1994); G. Kauffmann and S. Charlot, “Constraint on models of galaxy formation from the evolution of damped Lyman alpha absorption systems”, Berkeley preprint (1994); A. Klypin, S. Borgani, J. Holtzman and J. Primack, Perugia preprint (1994).
- [19] M. Davis, F. J. Summers and D. Schlegel, *Nature* **359** (1992) 393.
- [20] D. Yu. Pogosyan and A. A. Starobinsky, *Mon. Not. Roy. astr. Soc.* **265** (1993) 507.
- [21] R. K. Schaefer and Q. Shafi, *Phys. Rev. D* **49** (1994) 4990.
- [22] S. A. Bonometto, S. Borgani, S. Ghigna, A. Klypin and J. R. Primack, submitted to *Mon. Not. Roy. astr. Soc.* (1993).
- [23] G. Dvali, Q. Shafi and R. Schaefer, “Large Scale Structure and Supersymmetric Inflation Without Fine Tuning”, Pisa preprint (1994).
- [24] L. A. Kofman, N. Y. Gnedin and N. A. Bahcall *Astrophys. J.* **413** (1993) 1; R. Cen, N. Y. Gnedin and J. P. Ostriker, *Astrophys. J.* **417** (1993) 387; R. Cen and J. P. Ostriker, “X-ray clusters in a CDM+ $\Lambda$  universe; a direct, large scale, high resolution, hydrodynamic simulation” Princeton preprint (1994); E. F. Bunn and N. Sugiyama, “Cosmological-Constant Cold Dark Matter Models and the COBE Two-Year Sky Maps”, Berkeley preprint (1994).
- [25] N. Sugiyama and J. Silk, “The imprint of  $\Omega$  on the cosmic microwave background”, to appear in *Phys. Rev. Lett.* (1994).
- [26] M. Tegmark and J. Silk, “Reionization in an open CDM universe: implications for cosmic microwave background fluctuations”, Berkeley preprint (1994).
- [27] M. Kamionkowski, D. N. Spergel and N. Sugiyama, *Astrophys. J.* **426** (1994) L57; M. Kamionkowski and D. N. Spergel, to appear in *Astrophys. J.* (1 September 1994); B. Ratra and P. J. E. Peebles, “CDM cosmogony in an open universe”, Princeton preprint (1994); M. Kamionkowski, B. Ratra, D. N. Spergel and N. Sugiyama, “CBR anisotropy in an open inflation, CDM cosmogony”, Princeton preprint (1994); B. Ratra, ‘CDM cosmogony in an open universe’, Princeton preprint (1994); P. Coles and G. Ellis, “The case for an open universe”, QMW preprint (1994).
- [28] E. M. Lifshitz, *J. Phys. (Moscow)* **10** (1946) 116.
- [29] P. J. E. Peebles, *The Large Scale Structure of the Universe* (Princeton University Press, 1980).
- [30] G. Efstathiou, in “The Physics of the Early Universe”, eds Heavens, A., Peacock, J. and Davies, A. (SUSSP publications, 1990); C.-P. Ma and E. Bertschinger, “Cosmological Perturbation Theory in the Synchronous vs. Conformal Newtonian Gauge”, MIT preprint (1994).
- [31] J. M. Bardeen, *Phys. Rev. D* **22** (1980) 1882.
- [32] V.F. Mukhanov, H. A. Feldman and R. H. Brandenberger, *Phys. Rep.* **215** (1992) 203; H. Kodama and M. Sasaki, *Prog. Theor. Phys.* **78** (1984) 1; N. Sugiyama, N. Gouda and M. Sasaki, *Astrophys. J.* **365** (1990) 432 and references cited there; R. K. Schaefer, *Int. J. Mod. Phys. A* **6** (1991) 2075; W. Hu and N. Sugiyama, “Anisotropies in the cosmic microwave background”, Berkeley preprint (1994).

- [33] S. W. Hawking, *Astrophys. J.* **145** (1966) 544; D. W. Olson, *Phys. Rev. D* **14** (1976) 327.
- [34] G. F. R. Ellis and M. Bruni, *Phys. Rev. D* **40** (1989) 1804; M. Bruni, P. K. S. Dunsby and G. F. R. Ellis, *Astrophys. J.* **395** (1992) 34; P. K. S. Dunsby, M. Bruni and G. F. R. Ellis, *Astrophys. J.* **395** (1992) 54.
- [35] D. H. Lyth, *Phys. Rev. D* **31** (1985) 1792 .
- [36] M. Bruni and D. H. Lyth, *Phys. Letts. B* 323 (1994) 118.
- [37] D. H. Lyth and M. Mukherjee, *Phys. Rev. D* **38** (1988) 485.
- [38] D. H. Lyth and E. D. Stewart, *Astrophys. J.* **361** (1990) 343.
- [39] D. H. Lyth and E. D. Stewart, *Phys Lett.* **B252**, 336 (1993).
- [40] E. J. Copeland, A. R. Liddle, D. H. Lyth, E. D. Stewart and D. Wands, *Phys. Rev. D* **49** (1994) 6410; E. J. Copeland *et al*, this volume; E. D. Stewart, “Inflation, Supergravity and Superstrings”, Kyoto preprint (1994); E. D. Stewart, “Mutated Hybrid Inflation”, Kyoto preprint (1994).
- [41] S. Mollerach, S. Matarrese and F. Lucchin, “Blue Perturbation Spectra from Inflation”, CERN preprint (1993).
- [42] R. Harrison, *Phys. Rev. D* **1** (1970) 2726. Ya. B. Zel’dovich, *Astron. Astrophys.* **5** (1970) 84.
- [43] A. A. Starobinsky, *Phys. Lett.* **B117** (1982) 175; S. W. Hawking, *Phys. Lett.* **B115** (1982) 339; A. H. Guth and S.-Y. Pi, *Phys. Rev. Lett.* **49** (1982) 1110; J. M. Bardeen, P. S. Steinhardt and M. S. Turner, *Phys. Rev. D* **28** (1983) 679.
- [44] V. F. Mukhanov, *JETP Lett.* **41** (1985) 493 .
- [45] M. Sasaki, *Prog. Theor. Phys.* **76** (1986) 1036.
- [46] Stewart, E. D., Lyth, D. H., *Phys. Lett.* **B302** (1993) 171.
- [47] T. J. Allen, B. Grinstein and M. B. Wise, *Phys. Lett.* **B197** (1987) 66; J. D. Barrow and P. Coles, *Mon. Not. R. ast. Soc.* **244** (1990) 188; I. Yi and E. T. Vishniac, *Phys. Rev. D* **48** (1993) 950; T. Falk, R. Rangarajan and M. Srednicki, *Astrophys. J.* **403** (1994) L1; A. Gangui, F. Lucchin, S. Matarrese and S. Mollerach, *Astrophys. J.* (1994) to appear; A. Gangui, preprint (1994); K. Yamamoto and M. Sasaki, preprint (1994).
- [48] A. A. Starobinsky, *Sov. Astron. Lett.* **11** (1985) 133.
- [49] D. Scott, M. Srednicki and M. White, *Astrophys. J. Lett.* **421** (1994) L5.
- [50] R. Crittenden, J. R. Bond, R. L. Davis, G. Efstathiou, P. J. Steinhardt and M. S. Turner, *Phys. Rev. Lett.* **71** (1993) 324.
- [51] P. J. E. Peebles, *Astrophys. J. Lett.* **263** (1982) L1.
- [52] J. A. Peacock and S. J. Dodds, *Mon. Not. R. astr. Soc.* **267**, (1994) 1020.
- [53] A. Dekel, “Dynamics of Large-Scale Motions in the Universe”, to appear in *Ann. Rev. Astron. Astroph.*, October (1994).
- [54] A. Dekel *et al*, *Astrophys. J.* **412** (1993) 1.

- [55] C. Lacey and S. Cole, “Merger Rates in Hierarchical Models of Galaxy Formation. II. Comparison with N-Body Simulations”, Oxford preprint (1994).
- [56] C.-P. Ma and E. Bertschinger, “Do Galactic Systems Form Too Late in Cold+Hot Dark Matter Models?”, preprint (1994).
- [57] S. D. M. White, G. Efstathiou and C. S. Frenk, *Mon. Not. Roy. astr. Soc.* **262** (1993) 1023.
- [58] R. G. Carlberg *et al*, ‘Mapping moderate redshift clusters’, preprint (1993).
- [59] M. G. Haehnelt, *Mon. Not. Roy. astr. Soc.* **265** (1993) 727.
- [60] A. M. Wolfe, *Annals of the New York Academy of Sciences* **688** (1993) 281.
- [61] W. H. Zurek *et al*, in *Proc. of the 9th IAP Meeting*, eds. F. R. Bouchet and M. Lachieze-Rey (Editions Frontieres, gif-sur-Yvette, 1993), p. 465; J. G. Bartlett and A. Blanchard, *ibid*, p. 281.
- [62] H. J. Mo, Y. P. Jing and G. Börner, *Mon. Not. Roy. astr. Soc.* **264** (1993) 825.
- [63] G. B. Dalton *et al*, “The two-point correlation function of rich clusters of galaxies”, Oxford preprint (1994).
- [64] K. Freese, J. A. Frieman and A. V. Olinto, *Phys. Rev. Lett.* **65** (1990) 3233.
- [65] A. R. Liddle, D. H. Lyth, R. K. Schaefer, Q. Shafi and P. Viana, in preparation (1994).
- [66] D. P. Schneider, M. Schmidt and J. E. Gunn, *Astron. J.* **98** (1989) 1507.
- [67] M. Tegmark, J. Silk and A. Blanchard, *Astrophys. J.* **420** (1994) 484.

## Figure Captions

*Figure 1. [NOT AVAILABLE VIA BULLETIN BOARD]*

This figure is reproduced from [50], and it shows predictions for the spectrum  $\mathcal{P}_T(l)$  of the cmb anisotropy. The full line ‘S’ shows the contribution (normalized to 1 at  $l = 2$ ) of an adiabatic density perturbation with  $n = 0.85$ ,  $h = 0.5$ ,  $\Omega_B = 0.05$  and critical total density. It was calculated for pure CDM, but it is insensitive to the nature of the dark matter. (The dashed line is the prediction with  $\Omega_B = 0.01$  and the same value for  $h$ , but note that these quantities are actually linked by the nucleosynthesis relation  $\Omega_B h^2 = 0.013 \pm .002$ .) The line ‘T’ shows the gravitational wave contribution, also normalized to 1 at  $l = 2$ . As discussed in the text, it is not yet known whether this contribution is actually significant.

*Figures 2a and 2b.*

Observational data, shown in comparison to the canonical CDM model plus four variants, two with a hot component, one tilted and one with a low Hubble constant. All of the models are normalized on large scales to the COBE data as described in the text, and this data is represented schematically by the filled triangle at  $Rh^{-1} \sim 4000 \text{ Mpc}$ . The rest of the data comprises damped Lyman alpha systems and quasars (two lower limits on the left hand side), cluster abundance (triangle, with a tilted error bar indicating the dependence of the prediction on the assumed cluster mass), galaxy correlations (squares, error bars explained in text and in figure 3 caption) and bulk velocities (star). The second figure is a closeup of the first, which omits the COBE point.



*Figure 3.*

This is the same as Figures 2a and 2b except that everything has been divided by the prediction of the canonical CDM model. As explained in the text the galaxy correlation points enclosed by the ‘error box’ can be moved up and down more or less independently within the box, and in addition the whole box can be moved up and down by the amount indicated by the outer error bars. With this exception all of the error bars are more or less statistically independent, and represent something like a  $1\text{-}\sigma$  confidence level. The lower bounds are intended to express a high level of confidence as explained in the text.





